INDEXING AND ABSTRACTING

ABSTRACTING

Course Outline

Part I Abstracting

- Definition of Abstract
- Purposes for Abstracts
- Types of Abstracts
- Parts of an Abstract
- Why Write an Abstract
- Some Qualities of a Good Abstract
- Bibliographic Citation or Identification
- Processes for Writing Abstracts
- How to Make an Abstract
- Key Issues in Preparing Abstracts
- Steps for Writing Effective Abstracts
- Some Do's Don'ts for Abstractors
- Abstracts in Specific Disciplines

Part II Indexing

- What is Indexing?
- Who Does Indexing?
- How is Indexing Done?
- Indexes in History
- Skills or Education of Indexing
- Methodologies in Indexing
- Indexing Systems
- Indexing Languages
- Levels of Indexing
- Procedures in Indexing
- Principles of Indexing
- Rules in Indexing
- Indexes and Indexers
- Indexing Evaluation Checklist
- How Information Retrieval Started
- Thesaurus
- Thesaurus Management Software
- The NLM MeSH

Introduction: Writing Abstracts

- Abstracts are formal summaries writers prepare of their completed work. Abstracts are important tools for readers, especially as they try to keep up with an explosion of information in print and on the Internet.

Definition of Abstract

Abstracts, like all summaries, cover the main points of a piece of writing. Unlike executive summaries written for non-specialist audiences, abstracts use the same level of technical language and expertise found in the article itself. And unlike general summaries which can be adapted in many ways to meet various readers' and writers' needs, abstracts are typically 150 to 250 words and follow set patterns.

Because readers use abstracts for set purposes, these purposes further define abstracts.

Purposes for Abstracts

Abstracts typically serve five main goals:

- Help readers decide if they should read an entire article

  Readers use abstracts to see if a piece of writing interests them or relates to a topic they're working on. Rather than tracking down hundreds of articles, readers rely on abstracts to decide quickly if an article is pertinent. Equally important, readers use abstracts to help them gauge the sophistication or complexity of a piece of writing. If the abstract is too technical or too simplistic, readers know that the article will also be too technical or too simplistic.

- Help readers and researchers remember key findings on a topic

  Even after reading an article, readers often keep abstracts to remind them of which sources support conclusions. Because abstracts include complete bibliographic citations, they are helpful when readers begin writing up their research and citing sources.

- Help readers understand a text by acting as a pre-reading outline of key points

  Like other pre-reading strategies, reading an abstract before reading an article helps readers anticipate what's coming in the text itself. Using an abstract to get an overview of the text makes reading the text easier and more efficient.

- Index articles for quick recovery and cross-referencing

Even before computers made indexing easier, abstracts helped librarians and researchers find information more easily. With so many indexes now available electronically, abstracts with their keywords are even more important because readers can review hundreds of abstracts quickly to find the ones most useful for their research. Moreover, cross-referencing through abstracts opens up new areas of research that readers might not have known about when they started researching a topic.

- Allow supervisors to review technical work without becoming bogged down in details

  Although many managers and supervisors will prefer the less technical executive summary, some managers need to keep abreast of technical work. Research shows that only 15% of managers read the complete text of reports or articles. Most managers, then, rely on the executive summary or abstract as the clearest overview of employees' work.

Types of Abstracts

Although you'll see two types of abstracts—informative and descriptive—most writers now provide informative abstracts of their work. Click below to learn more about these two types of abstracts:

- Descriptive Abstract
- Informative Abstract
- A More Detailed Comparison of Descriptive vs. Informative

Descriptive Abstract

A descriptive abstract outlines the topics covered in a piece of writing so the reader can decide whether to read the entire document. In many ways, the descriptive abstract is like a table of contents in paragraph form. Unlike reading an informative abstract, reading a descriptive abstract cannot substitute for reading the document because it does not capture the content of the piece. Nor does a descriptive abstract fulfill the other main goals of abstracts as well as informative abstracts do. For all these reasons, descriptive abstracts are less and less common. Check with your instructor or the editor of the journal to which you are submitting a paper for details on the appropriate type of abstract for your audience.

Sample Descriptive Abstract

"Bonanza Creek LTER [Long Term Ecological Research] 1997 Annual Progress Report"
http://www.lter.alaska.edu/pubs/1997pr.html

We continue to document all major climatic variables in the uplands and floodplains at Bonanza Creek. In addition, we have documented the successional changes in microclimate in 9 successional upland and floodplain stands at Bonanza Creek (BNZ) and in four elevational locations at Caribou-Poker Creek (CPCRW). A sun photometer is operated cooperatively with NASA to estimate high-latitude atmospheric extinction coefficients for remote-sensing images. Electronic data are collected monthly and loaded into a database which produces monthly summaries. The data are checked for errors, documented, and placed on-line on the BNZ Web page. Climate data for the entire state have been summarized for the period of station records and krieged to produce maps of climate zones for Alaska based on growing-season and annual temperature and precipitation.

Informative Abstract

An informative abstract provides detail about the substance of a piece of writing because readers will sometimes rely on the abstract alone for information. Informative abstracts typically follow this format:

1. Identifying information (bibliographic citation or other identification of the document)
2. Concise restatement of the main point, including the initial problem or other background
3. Methodology (for experimental work) and key findings
4. Major conclusions

Informative abstracts usually appear in indexes like *Dissertation Abstracts International*; however, your instructor may ask you to write one as a cover sheet to a paper as well.

Sample Informative Abstract based on Non-experimental Work

Environmental Impact Statement. Federal Register: December 11, 1997 (Volume 62, Number 238). "Endangered and Threatened Wildlife and Plants; Proposed Revision of Special Regulations for the Gray Wolf." Fish and Wildlife Service, Department of the Interior.

http://www.epa.gov/fedrgstr/EPA-SPECIES/1997/December/Day-11/e32440.htm

On November 22, 1994, the U.S. Fish and Wildlife Service published special rules to establish nonessential experimental populations of gray wolves (Canis lupus) in Yellowstone National Park and central Idaho. The nonessential experimental population areas include all of Wyoming, most of Idaho, and much of central and southern Montana. A close reading of the special regulations indicates that, unintentionally, the language reads as though wolf control measures apply only outside of the experimental population area. This proposed revision is intended to amend language in the special regulations so

that it clearly applies within the Yellowstone nonessential experimental population area and the central Idaho nonessential experimental population area. This proposed change will not affect any of the assumptions and earlier analysis made in the environmental impact statement or other portions of the special rules. (accessed January 26, 1998)

Sample Informative Abstract based on Experimental Work

Palmquist, M., & Young, R. (1992). The Notion of Giftedness and Student Expectations About Writing. *Written Communication,* 9(1), 137-168.

Research reported by Daly, Miller, and their colleagues suggests that writing apprehension is related to a number of factors we do not yet fully understand. This study suggests that included among those factors should be the belief that writing ability is a gift. Giftedness, as it is referred to in the study, is roughly equivalent to the Romantic notion of original genius. Results from a survey of 247 postsecondary students enrolled in introductory writing courses at two institutions indicate that higher levels of belief in giftedness are correlated with higher levels of writing apprehension, lower self-assessments of writing ability, lower levels of confidence in achieving proficiency in certain writing activities and genres, and lower self-assessments of prior experience with writing instructors. Significant differences in levels of belief in giftedness were also found among students who differed in their perceptions of the most important purpose for writing, with students who identified "to express your own feelings about something" as the most important purpose for writing having the highest mean level of belief in giftedness. Although the validity of the notion that writing ability is a special gift is not directly addressed, the results suggest that belief in giftedness may have deleterious effects on student writers.

A More Detailed Comparison of Descriptive vs. Informative

The typical distinction between descriptive and informative is that the descriptive abstract is like a table of contents whereas the informative abstract lays out the content of the document. To show the differences as clearly as possible, we compare a shortened Table of Contents for a 100-page legal argument presented by the FDA and an informative abstract of the judge's decision in the case.

- Descriptive Table of Contents of the Argument
- Informative Abstract of the Decision

Table of Contents of the Argument

Court Brief (edited Table of Contents) Filed Dec. 2, 1996, by the Department of Justice in defense of FDA's determination of jurisdiction over cigarettes and smokeless tobacco

products and its regulations restricting those products to protect children and adolescents. http://www.usdoj.gov/civil/cases/tocnts.htm **Statement of the matter before the court; statement of material facts**

I. The health effects of cigarettes and smokeless tobacco
II. The basis for the assertion of jurisdiction
    A. The evidence that nicotine in cigarettes and smokeless tobacco "affect[s] the structure or any function of the body"
    B. The evidence that the pharmacological effects of nicotine in cigarettes and smokeless tobacco are "intended"
    C. The evidence that cigarettes and smokeless tobacco are "combination products"

III. The rule
    A. Cigarettes and smokeless tobacco as combination products
    B. The regulatory goal
    C. Youth access restrictions
    D. Advertising and promotion restrictions

Questions Presented

Arguments

Congress has not precluded FDA from regulating cigarettes and smokeless tobacco under the FDCA.

I. "Customarily marketed" cigarettes and smokeless tobacco are not exempt from regulation under the FDCA
    A. Standard of review: Chevron, U.S.A., Inc. v. Natural Resources Defense Council, Inc.
    B. Chevron, step one
    C. Chevron, step two: FDA's application of the FDCA to cigarettes and smokeless tobacco is "based on a permissible construction of the statute"

II. The Federal Cigarette Labeling and Advertising Act, Comprehensive Smokeless Tobacco Health Education Act, and the Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act do not foreclose FDA from regulating cigarettes and smokeless tobacco under the FDCA
    A. No statute, or combination of statutes, can override the FDCA in the absence of express preclusion or other clearly expressed Congressional intent
    B. Federal Cigarette Labeling and Advertising Act
    C. Comprehensive Smokeless Tobacco Health Education Act
    D. Alcohol, Drug Abuse, and Mental Health Administration Reorganization Act

III. The separation of powers doctrine does not prohibit FDA's regulation of tobacco products

Nicotine in cigarettes and smokeless tobacco is a drug, and cigarettes and smokeless tobacco are drug delivery devices under the FDCA.

I. Cigarettes and smokeless tobacco fall squarely within the Act's drug and device definitions
   A. Cigarettes and smokeless tobacco "affect the structure or any function of the body"
   B. Nicotine's effects are intended by the manufacturers

II. FDA's application of the medical device provisions to cigarettes and smokeless tobacco does not affect FDA's jurisdiction over these products
   A. Cigarettes and smokeless tobacco are combination drug/device products and may be regulated under the Act's device authorities
   B. FDA's application of device provisions to cigarettes and smokeless tobacco is reasonable

The restrictions imposed by FDA on advertising and other promotion of cigarettes and smokeless tobacco are fully consistent with the first amendment.

I. The agency's regulations must be judged pursuant to the Supreme Court's Central Hudson standard
   A. The Central Hudson standard and the proper First Amendment analysis
   B. Recent rulings by the Supreme Court in 44 Liquormart, and by the Fourth Circuit in Anheuser-Busch and Penn Advertising
   C. In applying the Central Hudson test, the Court's decision should be based on the record created by the Agency, and the reasonable determinations made by FDA are not to be disregarded

II. The government's interest here is plainly substantial
III. FDA has demonstrated that advertising affects tobacco use by minors, to the detriment of the public health, and that the agency's restrictions on advertising of these products should alleviate that problem to a material degree
IV. FDA's advertising restrictions are narrowly tailored
   A. The restrictions are designed to preserve the flow of information to lawful consumers
   B. The availability of non-Speech related regulatory alternatives does not invalidate FDA's regulations
   C. Each of FDA's individual advertising restrictions is narrowly tailored

Conclusion

Informative Abstract of the Decision

Summary of Federal District Court's Ruling on FDA's Jurisdiction Over, and Regulation of, Cigarettes and Smokeless Tobacco May 2, 1997
http://www.fda.gov/opacom/backgrounders/bg97-9.html

On April 25, 1997, Judge William Osteen of the Federal District Court in Greensboro, North Carolina, ruled that FDA has jurisdiction under the Federal Food, Drug, and Cosmetic Act to regulate nicotine-containing cigarettes and smokeless tobacco. The Court held that "tobacco products fit within the FDCA's definitions of 'drug' and 'device,'" and that FDA can regulate cigarettes and smokeless tobacco products as drug delivery devices under the combination product and restricted device provisions of the Act.

With respect to the tobacco rule, the Court upheld all restrictions involving youth access and labeling, including two access provisions that went into effect Feb. 28: (1) the prohibition on sales of cigarettes and smokeless tobacco products to children and adolescents under 18, and (2) the requirement that retailers check photo identification of customers who are under 27 years of age.

The Court also upheld additional access and labeling restrictions originally scheduled to go into effect Aug. 28, 1997, including a prohibition on self-service displays and the placement of vending machines where children have access to them. The Court also upheld the ban on distribution of free samples, the sale of so-called kiddie packs of less than 20 cigarettes, and the sale of individual cigarettes. However, the Court delayed implementation of the provisions that have not yet gone into effect pending further action by the Court.

The Court invalidated on statutory grounds FDA's restrictions on the advertising and promotion of cigarettes and smokeless tobacco. Judge Osteen found that the statutory provision relied on by FDA, section 520(e) of the Act (21 U.S.C. 360j(e)), does not provide FDA with authority to regulate the advertising and promotion of tobacco products. Specifically, the Court found that the authority in that section to set "such other conditions" on the sale, distribution, or use of a restricted device does not encompass advertising restrictions. Because Judge Osteen based his ruling on the advertising provisions on purely statutory grounds, he declined to consider the First Amendment challenge to those parts of the rule. The government is appealing the advertising portion of the ruling.

(accessed January 26, 1998)

Bibliographic Citation or Identification

As more and more databases are stored and accessed electronically, abstracts are more frequently reproduced apart from the entire article or document. In a large corporation or

government entity, for instance, an abstract of a progress report might be circulated and stored in a dozen offices or on multiple computers even though the report itself is filed in only one location. Clear identification is crucial so that readers who want to review the entire text can locate it from the information given with the abstract.

Depending on where your writing is printed and stored, you'll need to include different kinds of identifying information with your abstract:

Bibliographic Citation

If your writing will be printed and disseminated as a book, part of a book, or an article in a journal or magazine, give a full bibliographic citation that includes all the publication information so that readers can find print copies of the article (even if your abstract will appear in unrelated electronic databases). For example, an abstract for a journal article begins with this citation:

Harris, L.D., & Wambeam, C.A. (1996). The Internet-Based Composition Classroom: A Study in Pedagogy. *Computers and Composition*, 13(3), 353-372.

Organizational Identification

If your abstract is part of a corporate or government document that will not be printed or disseminated outside the organization, you need only include your name, the title of the document, its completion date, a project name (if you produced the document as part of the work on a larger project), and an authorization or organizational number (if there is one).

If your abstract will be circulated outside your organization (for instance, if you work for a consulting company that writes reports for other companies), add to the information above: your company or organization name, the name of the organization that commissioned the document, a contract number (if there is one), a security classification (as appropriate for government documents), and key words to help in cataloguing your abstract.

Internet Citation

If you're "publishing" your own work on the World Wide Web or if your writing will appear on the Internet as part of a full-text electronic database, you can save readers time by citing the Internet address for the full text. Typically, writers note **both** print publication information and the URL (universal resource locator)--the *http* or *www* address--with the abstract.

For example, one of the abstracts cited in this module has this citation that includes both bibliographic information and the Internet address:

Environmental Impact Statement. "Endangered and Threatened Wildlife and Plants; Proposed Revision of Special Regulations for the Gray Wolf." Federal Register: December 11, 1997 (Volume 62, Number 238). Fish and Wildlife Service, Department of the Interior. *http://www.epa.gov/fedrgstr/EPA-SPECIES/1997/December/Day-11/e32440.htm*

Processes for Writing Abstracts

Unless you work for an abstracting service, you'll usually write abstracts of your own finished work. This section explores some strategies for drafting your abstract. Strategies and advice on revising and editing are located in **Key Issues in Preparing Abstracts.** Choose a method below to see which would best suit your writing process:

- Cut and Paste Method
- Outlining Method

Cut and Paste Method

Beginning with reading may seem odd since you wrote the paper, but it can frequently be the fastest way to write an abstract because it allows you to "lift" as much of the abstract from your original paper as possible.

- Guidelines for Using the Cut and Paste Method

1. As you read through your own paper, highlight or copy sentences which summarize the entire paper or individual sections or sub-points of your main argument.
2. Write (or copy) a sentence that summarizes the main point.
3. Add sentences that summarize sections (or write new sentences for sections that lack a concise summary sentence).
4. If you're writing a descriptive abstract, you're ready to begin revising.
5. If you're writing an informative abstract, look through your paper for details, particularly of key findings or major supporting arguments and major conclusions. Paste these into your abstract and proceed to editing for consistency and length--frequently in the original "cuts" you will still have more detail than is necessary in an abstract.

Outlining Method

Frequently, the best place to start writing an abstract is to first make an outline of the paper to serve as a rough draft of your abstract. The most efficient way to do this is to write what Kenneth Bruffee calls a descriptive or "backwards" outline.

Backwards Outline Instructions

1. Read through each paragraph of your paper and write one phrase or sentence that answers the question "what does this paragraph do?"
2. Take your list of descriptions for each paragraph and look for connections: i.e., do these 3 or 5 paragraphs do something similar? What is it?
3. When you've reduced your outline to 4 or 5 accurate generalizations, you most likely have a descriptive abstract.
4. If you're writing an informative abstract, fill in key details about your content.

Detailed Backwards Outline

Because informative abstracts need more detail, the regular backwards outline may not be as useful a strategy for this type of abstract. Instead, do a backwards outline on the left-hand side of a piece of paper. Then, on the right-hand side, answer the question "what does this paragraph *say?"* for each paragraph in the paper. Then complete the steps below:

1. Take your first column and generalize down to 4-5 sentences about what the paper *does*.
2. Use these sentences as topic sentences for the paragraphs in your abstract.
3. Now, go to your second column and choose appropriate content for each section you outlined in #2. In other words, use the right-hand column to fill in details about what your paper says on each point outlined in #2.

Key Issues in Preparing Abstracts

These are some of the key issues in writing an abstract:

Concise, Accurate Statement of the Main Idea

Abstracts begin with a one-sentence summary of the main point of your paper and often introduce the problem the paper explores. Especially for papers based on research, the first sentence (or two) of the abstract announces the subject and scope of the research as well as the problem and your thesis. That's quite a bit of information to condense into a

sentence or two, and so the concise statement of the main idea often takes careful revision.

Choose from these options to read more about preparing a concise statement of the main idea:

Condensing Information for Non-research Papers

Most non-research papers can be summed up in a nutshell statement—a single sentence that boils down a paper to its essential main point and doesn't aim to capture details, supporting arguments, or types of proof.

One-sentence Summaries for Different Types of Papers
Each of these non-research papers summarizes its main point based on its overall purpose:

This paper argues that the "saving democracy" rhetoric surrounding the Gulf War was merely a mask for the U.S.'s interest in keeping oil prices down. (From a political science paper whose purpose was to construct an *argument*.)

Ethnography and ethnology are the preferred research methods of many anthropologists. (From an anthropology paper whose purpose was to *inform* others about a research methodology.)

Condensing Information for Research Papers

In addition to stating the main point of the paper, research-based papers often need to set up the context and scope of the research as well. Setting the context includes stating the subject of your work as well as the problem that prompted your research. You might also refer to major researchers who have already done work on your topic as a way of setting the context. Remember, too, that your abstract must always include the main point of your paper, so don't neglect that focus as you work on stating the problem and context.:

- Example 1

  In this example, note that the writer uses the names of key researchers to set the context and then focuses on what researchers don't yet know. After setting up the problem he's addressing in the research, the writer then announces the scope and focus of the paper in the second sentence:

> Research reported by Daly, Miller, and their colleagues suggests that writing apprehension is related to a number of factors we do not yet fully understand. This study suggests that included among those factors should be the belief that writing ability is a gift. . . .

- Example 2

  In this example, the writer announces the subject and scope of the research although he doesn't set context or suggest the problem that prompted the research. Depending on your ultimate goals for the abstract, you may be more successful with this approach that states the main point of your research paper even without setting context:

  > This report examines the changes in photosynthesis with an energy-producing carnivorous plants, specifically the Venus Fly Trap. (From a botany research report which involved original lab research.)

Beware of Focusing too Narrowly

No one who has ever written a concise restatement of a complex point will claim that the work was easy or straightforward. Usually, a writer needs to work back and forth between revising the restatement and re-reading the paper to be sure the main idea is stated accurately and clearly. Having worked so hard on that point, though, don't assume that you don't need to revise other parts of your abstract. In this example, the writer restates only the main point and dismisses key information from the 15-page document that should be included in the abstract.

Sample Abstract with Overly Narrow Focus

Community Right-to-Know Notice. *Federal Register*: January 23, 1998 (Volume 63, Number 15). "Phosphoric Acid; Toxic Chemical Release Reporting." Environmental Protection Agency (EPA).

http://www.epa.gov/fedrgstr/EPA-WASTE/1998/January/Day-23/f1644.htm

EPA is denying a petition to delete phosphoric acid from the reporting requirements under section 313 of the Emergency Planning and Community Right-to-Know Act of 1986 (EPCRA) and section 6607 of the Pollution Prevention Act of 1990 (PPA). This action is based on EPA's conclusion that phosphoric acid does not meet the deletion criteria of EPCRA section 313(d)(3). Specifically, EPA is denying this petition because EPA's review of the petition and available information resulted in the conclusion that phosphoric acid meets the listing criterion in EPCRA section 313(d)(2)(C) in that the phosphates that result from the neutralization of phosphoric acid may cause algal blooms. Algal blooms result in deoxygenation of the water and other effects that may ultimately

lead to a number of serious adverse effects on ecosystems, including fish kills and changes in the composition of animal and plant life.

Test Your Ability to Judge Conciseness

The biggest problem writers run into when beginning an abstract is providing enough accurate information to convey an article's main idea without providing more detail than is needed. To test your ability to judge conciseness, click "Start Test" below, read the detailed summary, and then judge sample restatements of the main idea.

A Detailed Summary

A summary of: Jaime O'Neill, No Allusions in the Classroom, *Newsweek*, September 23, 1985.

Author Jaime O'Neill's article, "No Allusions in the Classroom," emphasizes the communication problem between teachers and students due to the students' lack of basic knowledge. The author supports this assertion by using a combination of personal experience, evidence obtained from recent polls, other professors' opinions, and the results of an experiment he conducted in his own classroom. The experiment O'Neill conducted was an ungraded eighty-six question "general knowledge" test issued to students on the first day of classes. On this test, "most students answered incorrectly far more often than they answered correctly." Incorrect answers included fallacies such as: "Darwin invented gravity" and "Leningrad was in Jamaica." Compounding the problem, students don't ask questions. This means that their teachers assume they know things that they do not. O'Neill shows the scope of this problem by showing that, according to their teachers, this seems to be a typical problem across the United States. O'Neill feels that common knowledge in a society is essential to communicate. Without this common knowledge, learning is made much more difficult because teacher and student do not have a common body of knowledge from which to draw. The author shows the deterioration of common knowledge through poll results, personal experience, other teachers' opinions, and his own experiment's results.

Sample Restatements of Main Idea

To test your ability to find a balance between insufficient/inaccurate information and too much information, judge these sample restatements of the main idea of Jaime O'Neill's article.

- In "No Allusions in the Classroom," professor Jaime M. O'Neill discusses the frustrations he had with his students not understanding his allusions.
  Is this a concise statement of the article's main idea?
  - Yes
  - No
- In the article "No Allusions in the Classroom," a teacher of 15 years, Jaime M. O'Neill exposes the overall lack of basic knowledge of college students in areas of

history, geography, and politics.
Is this a concise statement of the article's main idea?
    - Yes
    - No
- In the article "No Allusions in the Classroom," the author, Jaime M. O'Neill relays his frustrations resulting from what he finds to be a lack of basic or common knowledge of the American student. He is a professor who has taught at several different colleges in the United States. Through his experiences in the classroom he noticed the fact that students were not catching on to his allusions. He states in the article that this miscommunication is causing a problem in the education process.
  Is this a concise statement of the article's main idea?
    - Yes
    - No
- In "No Allusion in the Classroom," by Jaime M. O'Neill, the problem of ignorance is discussed. Jaime O'Neill says that college students lack common knowledge and it is interfering with education.
  Is this a concise statement of the article's main idea?
    - Yes
    - No
- Author Jaime O'Neill's article "No Allusions in the Classroom", emphasized the communication problem between teachers and students due to the students' lack of basic knowledge.
  Is this a concise statement of the article's main idea?
    - Yes
    - No

Organization of Subpoints

After a summary of the main topic/problem/point of your paper or report, the abstract provides some detail on how you reached this point. The information provided in the abstract should follow the organization of the paper/report itself, almost like providing an outline for the reader in text form.

- Abstracts of Papers Without Headings

  When abstracting a paper that doesn't have headings and sub-headings, you must depend on your sense of major "chunks" in the text. As you'll see in the following example, this writer followed his concise statement of the main point with two sentences that focus on the two main arguments presented in the paper.

  Note: *The numbers in this abstract are for illustration purposes only. Number 1 designates a concise statement of the main point. Number 2 designates a summary statement of the first major argument and its support (five pages in the original article). Number 3 corresponds to a summary of the second major*

argument (two pages in the original), and Number 4 corresponds to the second argument's support (two pages).

> **(1)** This paper argues that the "saving democracy" rhetoric surrounding the Gulf War was merely a mask for the U.S.'s interest in keeping oil prices down.**(2)** Such an argument is made by first describing the ways in which OPEC controlled oil prices by limiting sales, pointing specifically to how Kuwait was producing more oil than allowed by current OPEC agreements. **(3)** Second, the paper examines why the U.S. was invested in keeping good relations with the only two OPEC nations--Kuwait and Saudi Arabia--which frequently made trade agreements that benefited the U.S. **(4)** Finally, the paper does a close reading of the newspaper coverage of the Gulf War, examining how an early recognition of the monetary incentive changed to a democratic one when Bush ordered trOops to Saudi Arabia.

- Abstracts of Papers With Sub-Headings

  When abstracting a paper that has headings and sub-headings, use those to help you identify key parts of the paper for your abstract. The following sample abstract, based on a research paper, uses the introduction, subjects, methods, results, and discussion headings from the original paper.

  Note: The numbers in this abstract are for illustration purposes only. Number 1 designates a concise statement of the main point and "problem" prompting the research. Number 2 designates a summary of the selection of research subjects. Numbers 3 and 4 correspond to summaries of research methods and results, respectively, and Number 5 designates a summary of conclusions.

  > **(1)** "Students in networked classrooms" examines the question of whether students in a computer classroom are more likely to engage in peer review than students in traditional classrooms. **(2)** To test this question, two classes in each environment were studied. **(3)** An observer participated in all four classes for the duration of a semester, noting the nature of the interaction between students. Further, the observer interviewed both students and teachers about the nature of peer interaction and review. **(4)** Based on this sample, the study finds that students in computer classrooms are more likely, by a ratio of 2:1, to engage in peer review.**(5)** As a result of this finding, the paper concludes that for this one variable, computer classrooms are a more effective environment in which to teach writing.

Use of Details

Details should be used judiciously in abstracts. Determining the amount of detail to provide depends a great deal on what type of abstract you are writing (informative or

descriptive), the complexity of the paper, the word limit for the abstract, and the purposes you imagine readers of your abstract have for reading.

- Complexity of the Paper

  An abstract of a five-page progress report is likely to be shorter than an abstract for a 100-page Master's thesis, mainly because a long paper will include more main ideas, not just details. Keep in mind your readers and their reasons for reading your abstract. Focus your abstract on main ideas and provide only those details that are crucial for readers to understand your main points.

- Word Limit for the Abstract

  Some publications limit the length of abstracts to no more than 75 words. Others allow abstracts of complex documents to run up to 350 words. Be sure to check the publication's guidelines. If it has a low word limit, concentrate on capturing only main ideas from your paper. Don't try to cut a 200-word abstract down to 125 words by simply cutting connecting words, articles, etc. Even the shortest abstracts need to be readable, not telegraphic.

- Readers' Purposes

  Some publications limit the length of abstracts to no more than 75 words. Others allow abstracts of complex documents to run up to 350 words. Be sure to check the publication's guidelines. If it has a low word limit, concentrate on capturing only main ideas from your paper. Don't try to cut a 200-word abstract down to 125 words by simply cutting connecting words, articles, etc. Even the shortest abstracts need to be readable, not telegraphic.

Revising and Editing

When you work from your own texts, abstracts are usually easy to draft. After all, most writers begin by cutting and pasting from the text itself. But abstracts can be tricky to revise and edit, particularly if you need to reach a low word count. In this section, we offer some advice on strategies for moving from a first draft of an abstract to a polished finished version.

- Being Concise

  When you cut and paste parts of your paper into your draft abstract, you may find that you initially include words and phrases that clarify the meaning in the paper but that simply add extra words in the abstract. Read your drafts carefully to cut unnecessary words. Note that the italicized words in the example can be cut without any loss of meaning in the abstract.

> Palmquist, M. (1995). "Students in Networked Classrooms." *Computers and Composition, 10*(4), 25-57.
>
> "Students in networked classrooms" examines *the question of* whether students in a computer classroom are more likely to engage in peer review than students in traditional classrooms. *To test this question*, two classes in each environment were studied. An observer participated in all four classes for the *duration of a* semester, noting *the nature of the* interaction between students. Further, the observer interviewed both students and teachers about *the nature of* peer interaction and review. *Based on this sample*, the study finds that students in computer classrooms are more likely, by a ratio of 2:1, to engage in peer review. *As a result of this finding*, the paper concludes that, for this one variable, computer classrooms are a more effective environment in which to teach writing.

- Smoothing out Connections

  After you revise for conciseness, you will also want to be sure that each sentence in your abstract leads smoothly into the next. Sometimes you need to add or change transitional words and phrases. Sometimes you need to repeat key words. And sometimes, you need to combine sentences so that the connections between ideas are logically clear.

  In our example, we combine what were sentences 2 and 3 and the last two sentences.

  > Palmquist, M. (1995). "Students in Networked Classrooms." *Computers and Composition, 10*(4), 25-57.
  >
  > This paper examines whether students in a computer classroom are more likely to engage in peer review than students in a traditional classroom. Two classes in each environment were observed, with the participant-observer noting interactions between students. Further, the observer interviewed both students and teachers about peer interaction and review. The study finds that students in computer classrooms are twice as likely to engage in peer review and concludes that, for this one variable, computer classrooms are a more effective environment in which to teach writing.

- Avoiding Telegraphic Abstracts

  A highly condensed style can save money when you send a telegram but can make abstracts too dense. Don't cut articles (*a, an, the*) or connecting words that show relationships among ideas. Do repeat key words that show the content of your paper. Abstracts may be short, but they are meant to be readable.

- Polishing Style

A reader looks at a summary for the sole purpose of getting a quick glimpse of the article. As a result, she doesn't want to waste time with a lot of phrases and words that do not further the meaning, nor is she interested in the summary writer's opinion. Accounting for audience needs, there are three generalizable principles about the style of summaries:

- Use of "I"

  Although use of "I" or "we" is acceptable in some disciplines, many frown on its use in abstracts. Read several abstracts in the publication you're submitting to or the databases you expect to include your abstract. When in doubt, do not use "I." Instead, use the following strategies:

  Substitute for "I"
  Most abstracts make the paper/report/study the focus of the abstract and the grammatical subject of sentences in the abstract. Try these sentence openers:

  - This paper explores. . .
  - This study suggests. . .
  - The report investigates. . . .

  Passive Voice In combination with substitutes for "I," passive voice helps writers focus on the paper/report/study. Instead of, "I propose that ethnography is a better research method than case study" (active voice), the abstract might use: "Ethnography is proposed as a better research method than case study." (passive voice) Be sure to combine substitutes for "I" with passive voice to avoid overusing the passive.

- Use of Quotes

  When using your own sentences, you don't need to put them in quotation marks. For example, if your methods section begins with "Three methods were used to investigate this question: case study, surveys, and observational research," feel free to repeat the sentence in its entirety in the abstract. Remember, however, the following points:

  - Revise the sentence so it makes sense in the abstract (i.e., if you have not summarized "this question" in the abstract, omit substitute for that phrase).
  - Do not "lift" sentences which are not your own (i.e., quotes from other people's work).

- Use of Literary Present Tense

  Abstracts use the present tense because we assume texts speak to the present even if their authors are dead and/or wrote the words in the past. As a result, write about the text and/or author as if they were composing the words at the moment. For example:

  - Hemingway describes Paris as.....
  - The Declaration of Independence states that all men are created equal.

  Caution: This rule varies from discipline to discipline.

Abstracts in Specific Disciplines

Abstracts have common elements and uses, but read enough abstracts in your field to be aware of their specific details or differences. Choose from the examples below to see additional sample abstracts. The abstract from Civil Engineering includes instructor comments.

- Civil Engineering

  MASK Engineering has designed a performing arts center for the CSU campus in order to provide a complex that will better serve the campus and the community. This facility will not only improve the performing arts programs on campus but will encourage students and community members to attend more cultural events in Fort Collins. The capacity of the new facility will exceed that of existing structures on campus, and the quality of sound and aesthetics will be improved. Some of the features included are a large performing hall, a coffee shop, a banquet hall, and a recording studio. The total area of the complex is 56,500 square feet split into three levels.

  Instructor's Comments: This abstract summarizes the accomplishments of the project and what it will do. It also summarizes some of the actual design and indicates that it's going to include a performing hall, coffee shop, banquet hall, and recording studio.

  The writing, however, could be a little tighter in my opinion. The first sentence looks like it's around 20 words long. First of all, the expression "will better service the campus and the community" doesn't mean anything. What does "better serve" mean? A better choice might be, "MASK Engineering has designed a new Performing Arts Center that will meet the needs of the theater community," or something more specific.

  The second sentence is typical. It gives the particular vehicle for doing the programs. However, it implies that the facility improves programs, and I'm

not sure that's quite the right subject for this sentence. Furthermore, there's no point to the word "but" here. There's no contrast here, so this is a grammatical problem. This kind of problem can be avoided through careful reading, asking what each sentence accomplishes.

The abstract gets stronger after this. "The capacity of the new facility will exceed that" is very specific. "The quality, sound and aesthetics will be improved. Some of the features included are this..." The writers are very good at being descriptive. I think engineering students are more comfortable with the descriptive aspect of their material than with the lead-in.

- English

LeCourt, D. (1996. Composition's Theoretical Irony: WAC as Uncritical Pedagogy. *Journal of Advanced Composition, 16*(3), 389-406.

This paper argues that writing across the curriculum has failed to consider how its practices and theories serve to inscribe students within normalized discourses. As scholars such as Susan McLeod, Anne Herrington and Charles Moran begin to re-think the way writing-across-the-curriculum programs have situated themselves within composition theory, an intriguing disparity has presented itself between writing-to-learn and learning-to-write. As McLeod points out, these two approaches to WAC, which she designates the "cognitive" and the "rhetorical," respectively, exist in most programs simultaneously despite their radically different epistemological assumptions. This paper suggests, however, that despite the two approaches' seeming epistemological differences, they work toward a similar goal: the accommodation or inscription of (student) subjects into the various disciplinary strands of academic discourse. From a poststructural perspective, the goals of both these models function as a coherent technology of subject production. Writing to learn exercises provide a discursive space in which students learn to write themselves as subjects of the discourse, using the writing space to "practice" an integration of self with a disciplinary subjectivity. The rhetorical model reinforces such an integration even more strongly, providing explicit instruction in how the discursive subject must write herself in order to produce "effective" prose which mirrors the texts of other "speaking" subjects of the discourse. In sum, both approaches to WAC are subject to the same description and critique of how academic discourse seeks to inscribe students as subjects that has been forged against composition instruction in English departments (e.g., Schilb, Clifford, Faigley). Ironically, in WAC, we have presumed a clear mission for writing instruction that is not nearly so evident in our own approach to advanced literacy. The paper concludes, then, by offering yet a third model of WAC, one which suggests that students, as well as their instructors, engage in the investigative process of discovering how discursive conventions relate to their discipline's

epistemology and consider how that connection limits what can be said or thought within that discourse.

- Neurobiology

High Performance Computing Applications in Neurobiological Research; Muriel D. Ross, NASA Ames Research Center, Moffett Field, CA 94035; Kevin Montgomery, Sterling Software, Palo Alto, CA 94303; David G. Doshay, Sterling Software, Palo Alto, CA 94303; Thomas C. Chimento, Sterling Software, Palo Alto, CA 94303; Bruce R. Parnas, National Research Council Research Associate, Biocomputation Center, NASA Ames Research Center, Moffett Field, CA 94035

The human nervous system is a massively parallel processor of information. The vast numbers of neurons, synapses and circuits is daunting to those seeking to understand the neural basis of consciousness and intellect. Pervading obstacles are lack of knowledge of the detailed, three-dimensional (3-D) organization of even a simple neural system and the paucity of large scale, biologically relevant computer simulations. We use high performance graphics workstations and supercomputers to study the 3-D organization of gravity sensors as a prototypic architecture foreshadowing more complex systems. Scaled-down simulations run on a Silicon Graphics workstation and scaled-up, three-dimensional versions run on the Cray Y-MP and CM5 supercomputers.

To assist this research, we developed generalized computer-based methods for semiautomated, 3-D reconstruction of this tissue from transmission electron microscope (TEM) serial sections and for simulations of the reconstructed neurons and circuits. Sections are digitized directly from the TEM. Contours of objects are traced on the computer screen. Mosaicking images into sections, registration and visualization are automated. The same grids generated to connect contours for viewing objects provide tesselated surfaces for 1-D, 2-D and 3-D simulations of neuronal functioning. Finite element analysis of prism or segment volumes and color coding are used to track current spread after synapse activation. The biologically accurate simulation is reducible to a symbolic model that mimics the flow of information processing. Discharge patterns are displayed as spike trains. The symbolic model can be converted to an electronic circuit for potential implementation as a chip. The reconstructions can also be rendered in visual, sonic and tactile virtual media.

Using these methods, we demonstrated that gravity sensors are organized for parallel distributed processing of information. They have non-modular receptive fields that are organized into overlapping, dynamic cell assemblies. These provide a basis for functional degeneracy and graceful degradation. The sensors have two intrinsic microcircuits that are

prototypic of more advanced systems. These microcircuits are highly channeled (type I cell to a nerve terminal called a calyx) and distributed modifying (type II cells and feedforward/feedback neural lOops). A circuit of extrinsic origin likely biases the intrinsic circuits. We use simulation methods to study the effects of intrinsic feedback-feedforward lOops and of extrinsically driven biases on discharge patterns. These and similar investigations into the functioning of huge assemblies of neurons require supercomputer capabilities and pave the way for studies of human brain functioning as a grand challenge in supercomputer applications.

http://biocomp.arc.nasa.gov/papers/hpc_abstract.94.html

(accessed February 3, 1998)

- Geology

Department of the Interior - U.S. Geological Survey Inventory of Landslides Triggered by the 1994 Northridge, California Earthquake Edwin L. Harp and Randall W. Jibson Open-File Report 95-213 USGS Denver, CO 80225 1995

The 17 January 1994 Northridge, California, earthquake **(M=6.7)** triggered more than 11,000 landslides over an area of about 10,000 km². Most of the landslides were concentrated in a 1,000-km² area that includes the Santa Susana Mountains and the mountains north of the Santa Clara River valley. We mapped landslides triggered by the earthquake in the field and from 1:60,000-scale aerial photography provided by the U.S. Air Force and taken the morning of the earthquake; these were subsequently digitized and plotted in a GIS-based format, as shown on the accompanying maps (which also are accessible via Internet). Most of the triggered landslides were shallow (1-5 m), highly disrupted falls and slides in weakly cemented Tertiary to Pleistocene clastic sediment. Average volumes of these types of landslides were less than 1,000 m³, but many had volumes exceeding 100,000 m³. Many of the larger disrupted slides traveled more than 50 m, and a few moved as far as 200 m from the bases of steep parent slopes. Deeper (>5 m) rotational slumps and block slides numbered in the hundreds, a few of which exceeded 100,000 m³ in volume. The largest triggered landslide was a block slide having a volume of 8X10E06 m³. Triggered landslides damaged or destroyed dozens of homes, blocked roads, and damaged oil-field infrastructure. Analysis of landslide distribution with respect to variations in (1) landslide susceptibility and (2) strong shaking recorded by hundreds of instruments will form the basis of a seismic landslide hazard analysis of the Los Angeles area. *http://gldage.cr.usgs.gov/html_files/ofr95-213/ABSTRAC2.html* (accessed February 3, 1998)

A B S T R A C T I N G

Prof. Edelin A. Columnas
Faculty, MILS *(Master of Information and Library Systems)*
Department of Computer Application, School of Computer Studies, MSU-IIT, Iligan City

What is an Abstract?

An abstract is a condensed version of a longer piece of writing that highlights the major points covered, concisely describes the content and scope of the writing and reviews the writing's contents in abbreviated form. (LEO)

"Something that concentrates in itself the essential qualities of anything more extensive or more general, or of several things; essence." [*Urdang, 1973*]

Abstract, like all summaries, cover the main points of a piece of writing. Unlike executive summaries written for non-specialist audiences, abstracts use the same level of technical language and expertise found in the article itself. And unlike general summaries which can be adapted in many ways to meet various readers' and writers' needs, abstracts are typically 150 -250 words and follow set patterns.

It is not an extract, a summary, annotation, terse literature, precise, summary, resume, epitome, paraphrase, digest, or a synopsis. "It is like the juice of a lemon."

Parts of an Abstract's Content

The abstract is a very brief overview of your entire study. It tells the reader What the author did?. Why he did it?, How he did it?, What he found?, and What it means?. The abstract should briefly state the purpose of the research (introduction), how the problem was studied (methods), the principal findings (results), and what the findings mean (discussion and conclusion). It is important to be descriptive but concise –say only what essential, using no more words than necessary to convey meaning.

Table 1.0

a. Motivation (Purpose)	3 %
b. Problem Statement	7 %
c. Approach (Methodology)	15 %
d. Result (Discussion)	70 %
e. Conclusion/ Recommendation	5 %
	100 %

Parts continued…

Motivation:
Why do we care about the problem and the results? If the problem isn't obviously "interesting" it might be better to put motivation first; but if the work is incremental progress on a problem that is widely recognized as important, then it is probably better to put the problem statement first to indicate which piece of the larger problem are broken off. This section should include the importance of the work, the difficulty of the area, and the impact it might have if successful.

Problem statement:
What *problem* the work is trying to solve? What is the *scope* of the work (a generalized approach, or for a specific situation)? Be careful not to use too much jargon. In some cases it is appropriate to put the problem statement before the motivation, but usually this only works if most readers already understand why the problem is important.

Parts continued…

Approach:
How did the author go about solving or making progress on the problem? Did he use simulation, analytic models, prototype construction, or analysis of field data for an actual product? What was the *extent* of the work (did he look at one application program or a hundred programs in twenty different programming languages?) What important *variables* did he control, ignore, or measure?

Results:
What's the answer? Specifically, most good computer architecture papers conclude that something is so many percent faster, cheaper, smaller, or otherwise better than something else. Put the result there, in numbers. Avoid vague, hand-waving results such as "very", "small", or "significant."

Parts continued…

Conclusions:
What are the implications of the findings? Is it going to change the world (unlikely), be a significant "win", be a nice hack, or simply serve as a road sign indicating that this path is a waste of time (all of the previous results are useful). Are the results *general*, potentially generalizable, or specific to a particular case?

What Types of Abstracts are Typically Used?

1. Descriptive Abstracts

- Tell readers what information the report, article, or paper contains.
- Include the purpose, methods, and scope of the report, article, or paper.
- Do not provide results, conclusions, or recommendations.
- Are always very short, usually under 100 words.
- Introduced the subject to readers, who must then read the report, article, or paper to find out the author's results, conclusions, or recommendations.

2. Informative Abstract

- Communicate specific information from the report, article or paper.
- Include the purpose, methods, and scope of the report, article, or paper.
- Provide the report, article, or paper's results, conclusions, and recommendations.
- Are short – from a paragraph to a page or two, depending upon the length of the original work being abstracted. Usually informative abstracts are 10% or less of the length of the original piece.
- Allow readers to decide whether they want to read the report, article, or paper.

Why Write an Abstract?

1. The abstract is of utmost importance, for it is read by 10 to 500 times more people than hear or read the entire article.
2. Writing an abstract is ***hardwork,*** but will repay you with increased impact on the world by ***enticing*** people to read Info. is power.
3. Key words of an abstract is vital because of today's electronic information retrieval systems.
4. For current awareness services
5. For inquiry answering and desk research
6. For preparing briefing and writing reports
7. For database creation
8. For managers who need to assimilate documents rapidly and review them in reports. CEO's executive info system
9. The paradigm shift of the world from "collection to access"

Some Qualities of a Good Abstract?

Desired qualities	Reasons
Complete but concise description of the work.	Because on-line databases typically contain only abstracts. It should entice potential reader to obtain full text.
Fully self-contained, capsule descriptions of the original work.	A stand-alone document can be understood without reading the original work.
Follows a ***checklist*** of the method, i.e. uses an introduction/body/conclusion structure which presents the article, paper, or report's purpose, results, conclusions, and recommendations in that order.	To "sell your collection; to serve the need of business executives
Logic, clarity, development, and the scientific method are merged with condensation, compaction and summarization to repack multi-page manuscript into a single paragraph. [150 - 250 words or lesser]	For readability
Key facts and opinion are extracted from document and reproduced accurately	Because of Info. Overload.

How to Make an Abstract

Some Considerations: ***(Phil Koopman, Carnegie Mellon University 1997)***

- An abstract must be a fully self-contained, capsule description of the paper. It can't assume (or attempt to provoke) the reader into flipping through looking for an explanation of what is meant by some vague statement. It must make sense all by itself. Some points to consider include: Meet the word count limitation. If your abstract runs too long, either it will be rejected or someone will take a chainsaw to it to get it down to size. Your purposes will be better served by doing the difficult task of cutting yourself, rather than leaving it to someone else who might be more interested in meeting size restrictions than in representing your efforts in the best possible manner.

- Any major restrictions or limitations on the results should be stated, if only by using "weasel-words" such as "might", "could", "may", and "seem".

Some Considerations Continued…

- Think of a half-dozen search phrases and keywords that people looking for work might use. Be sure that those exact phrases appear in your abstract, so that they will turn up at the top of a search result listing.
- Some publications request "keywords". These have two purposes. They are used to facilitate keyword index searches, which are greatly reduced in importance now that on-line abstract text searching is commonly used. However, they are also used to assign papers to review committees or editors, which can be extremely important to your fate. So make sure that the keywords you pick are obvious, using chosen topic area as one of the keyword tuples).

Some Considerations Continued… ***(Kenneth A. Small, Dept. of Econ., Univ. of California at Irvine 1997)***

- The purpose is not to tell the reader you did something: it's to tell her what you did in the simplest, most informative way possible. Too many abstracts begin: ***"This article investigates the determinants of drug use."*** With just a few more words you can tell the reader what you are actually doing: ***"I use a model of rational addiction to show how institutions and tastes affect drug consumption."***
- Similarly for the concluding sentence. Instead of: ***"The results show that choice among drugs depends on the institutional framework and taste parameters,"*** how about: ***"The results show that the substitution of designer drugs for marijuana increases with media coverage of drug issues and with the propensity to schizophrenia."***
- Should you hold back your punch line to increase suspense? It's tempting, but an economics article is not a Sherlock Holmes story. Readers are too busy to appreciate the excitement of your little mystery tale: they want quick information. If it suits them, they will read on for the details, appreciating that you've already made it easier for them by explaining where you're headed.

Some Considerations Continued…

- The introduction and conclusion are the next most likely parts of the document to be read. Think of these sections as extensions of the abstract, carefully coordinated to entice the reader further and further into the details of the work.

- Abstracts for submission to meetings serve a somewhat different purpose, because the reader doesn't expect the results to be all worked out. Here it's more important to focus on the ***significance of the idea*** and ***how you will improve on existing knowledge***. So a few citations are warranted, enough to identify the literature to which the work is contributing and to explain how its differs from others. Explain what conclusions may emerge and what implications they would have.

Some Consideration Continued…(author…?)

- **Stress content not intent**
  An abstract should summarize the content of the article not the intent of the writer(s). It should contain only what is specifically reported in the manuscript. Based on Day's (1983) recommendations, the abstract should have at most one or two sentences on each of the four foundation points of your work. These are: (1) principal objectives and scope of the work, (2) methodology, (3) results, and (4) conclusions. Leave out extra baggage Too many abstracts are filled with extras. Stick to the four foundation points. Background references, literature surveys, and setting; justification and motivation; global definition of the problems and discipline; needs, benefits, claims, recommendations, opinions, and utility of the work are not part of the abstract. They may qualify as intent, motivation, bulwark, etc. and appear in the main body of the work, but they are not part of, nor are they needed in the abstract.

Some Considerations Continued…

- **Assume a knowledgeable reader**
  In all writing there is always the question of the knowledge level of the reader. In writing the abstract assume a knowledgeable reader. The level of knowledge assumed should be that of a good supervisor, one who understands the type of work, but is not active in that area and may not remember the more esoteric nomenclature. Abstracts require rethinking and reformulating.

- **Write the abstract last**
  After you read the document and let it ferment, you should be able to write a terse, but exact, description of each of the four foundation points based on the content of the written manuscript. If you cannot formulate a four-to-eight-sentence abstract, maybe you should reexamine the manuscript.

Some Considerations Continued…

- **Avoid passive voice**
  To avoid a conflict between intent, content, and extra baggage, it is very helpful to write in the active voice. Sentences structured with "is presented," "was done," "is given," "is developed," etc. are ***easy to write, terrible to read***, and convey essentially nothing to the reader. Despite how they sound, they provide no useful content; they show only intent, while giving no content.
- **Keep it short**
  As indicated above, your abstract does not need to be longer than an eight sentence maximum. If it is, it's either not doing its job and you have excess baggage or you have a very rare exception. I stress rare. You should complete your abstract in one paragraph only. It should not have multiple paragraphs. If it does, again, you probably have excess baggage. Because of this length restriction, your abstract may not be smooth and flowing. That's OK. It may be very terse, staccato-sounding, and not particularly easy to read. That's OK, too. Ease of reading is not a requirement; conveying content is. Because of its short length, a reader can easily read through it a second time.

Some Considerations Continued…

- **Make quantitative not qualitative statements**
  It is very annoying to read that A is bigger or smaller, faster or slower, longer or shorter, more or less porous or permeable, etc., than B without knowing how much or to what extent.

- **Don't use equations or other mathematical notation**
  It should be obvious that equations and mathematical notation have no place in an abstract, but some writers refuse to recognize this. Remember, the abstract must be able to stand alone. Give names to parameters not mathematical symbols. If required, name common equations (e.g., the wave equation, Laplace's equation, etc.) and summarize your contribution and findings from them.

Some Considerations Continued…

- **Empathize with the first-time reader**
  Your abstract is short and can be edited easily. As an editing tool, read your abstract and imagine that it is being read for the first time and that it is the only portion of the article that is read. Answer honestly, "what have I conveyed to the reader? With what does he/she walk away?" If, for example, it is a compendium of "is presents" and "was done," this is not an abstract.

  In the end, remember that your intent in writing an abstract is to showcase the work in a direct and lean manner. The abstract should give the casual reader a bit of useful information and the interested reader a prod to continue and read the article. It's a marketing tool, but not a platform for salesmanship or filibustering. Also remember that there are many services that copy and publish abstracts. An abstract must be structured to stand alone. It should not require leaning on the main body of the article.

Steps for Writing Effective Abstracts

1. ***Read or listen to the entire paper long enough for it to percolate and brew up a clear vision of what the paper is all about.***

2. ***Reread the article, paper, or report with the goal of abstracting in mind.***
   - Look specifically for these main parts of the article, paper, or report: purpose, methods scope, results, conclusions, and recommendation or look for the what (the author did), Why (the author did it?) What (the author found), What (it means).
   - Use the headings, outline heads, and table of contents as a guide to writing an abstract.
   - The introduction and the summary are good places to begin. These generally cover what the article emphasize.

Steps Continued…

3. ***After you've finished rereading article, paper, or report write a rough draft without looking back at what you're abstracting.***
   - Don't merely copy key sentences from the article, paper, or report: you'll put in too much or too little information.
   - Don't rely on the way materials was phrased in the article, paper, or report: summarize information in a new way.
4. ***Revise your rough draft to:***
   - correct weaknesses in organization
   - Improve transitions from point to point
   - Drop unnecessary information
   - Add important information you left out.
   - Eliminate wordiness.
   - Fix errors in grammar, spelling, and punctuation
   - Flavor it with style (avoid verbosity, use of active voice than passive voice, condense parts if possible, etc. (get tips from do's and don'ts)
5. ***Print your final copy and read it again to catch any glitches that you find.***

Some Requisites/ Tips in Abstracting

1. Writing and reading skills
2. Sign-posting and skim reading (summarizing a material)
3. Scanning documents that have little signposting (explaining what the book is about)
4. Skills in cutting through the verbiage (essentializing an article)
5. Tackling documents that offer few clues (rapid reading)
6. Making abstracts stylish and comprehensive

Some Do's and Don'ts for Abstractor

DO's

1. Scan the document purposefully for key facts.
2. Slant the abstract to the audience.
3. Tell what was found.
4. Tell why the work was done.
5. Tell how the work was done.
6. Place findings early in the topical sentence.
7. Put details in succeeding sentence.
8. Place general statements last.
9. Separate relatively independent subjects.
10. Differentiate experiment from hypothesis.
11. Be informative but brief.
12. Be exact, concise and unambiguous.
13. Use short, complete sentences.
14. Use short, simple familiar words.
15. Avoid unnecessary words.
16. Employ normal technical English.
17. Use direct statements.
18. Describe conclusions in the present tense.
19. Use abbreviations sparingly
20. Avoid using unusual words or phrases.
21. Cite bibliographic data completely.

Don'ts

1. Repeat words of the title
2. Change the meaning of the original
3. Comment on or interpret the document. Do not use adjectives
4. Mention earlier work.
5. Include detailed experimental results.
6. Discuss details of conventional apparatus.
7. Mention future work.
8. Begin abstract with stock phrases.
9. Use uncommon or rare phraseology.
10. Use questionable jargon.
11. Waste words by stating the obvious.
12. Say the same things in two days.
13. Use noun forms of verbs.
14. Overuse synonyms.
15. Use a choppy, telegraphic style.
16. It should not be a mere recital of subjects covered. Expressions such as "is discussed" and "is described" should never be included!
17. The abstract should be a condensation and concentration of the *essential information* in the paper.

Exercise 1.0 On Verbosity

1. Subjects of these study were
2. Owing to the fact that
3. During the period of time that
4. Square in shape
5. Stated verbally
6. Has been found to increase
7. It would seem that
8. Blue in Color
9. At the hour at noon
10. Entirely put into completion
11. Experience of an educational nature

Workshop 1.0 Abstracting a Thesis

Some projects may not lend themselves to this format, so don't feel that you need to use this worksheet.)

Use one or two concise sentences to summarize the most important aspects of the work listed below.

Scope (What is this project about? Why is this project interesting or important?)

Purpose (Intent/ Objective of the author)

Hypothesis (What did you think you would find? Why?)

Methods (Briefly explain your procedure.)

Results (What did the author find when you performs his experiment?)

Discussion (Are the results consistent with the initial hypothesis? Why or why not?)

Conclusion (What is *the author's* interpretation of what these results mean? Why should anyone become excited about or interested in the findings?)

Abstracting An Article

Workshop 2

Scrutiny of the abstract I by Kenneth K. Landes (University of Michigan, Ann Arbor, Michigan

Choice 1

A partial biography of the writer is given. The inadequate abstract is discussed. What should be covered by an abstract is considered. The importance of the abstract is described. Dictionary definitions "Abstract" are quoted. At the conclusion a revised abstract is presented.

Choice 2

The abstract is of utmost importance, for, for it is read by 10 to 500 times more people than hear or read the entire article. It should not be a mere recital of subjects covered. Expressions such as "is discussed" and "is described" should never be included! The abstract should be a condensation and concentration of the essential information in the paper.

Abstracting An Article

Workshop 3

Scrutiny of the abstract II

Choice 1

The abstract is of utmost importance for it is read by 10 to 500 times more people than hear or read the entire article. It should not be a mere recital of the subjects covered, replete with such expressions as "is discussed" and "is described." It should be condensation and concentration of the essential qualities of the paper.

Choice 2

The behavior of editor is discussed. What should be covered by an abstract is considered. The importance of the abstract is described. Dictionary definitions of "abstract" are quoted. At the conclusion a revised abstract is presented

Abstracting An Article

Workshop 3

Scrutiny of the abstract II

Choice 1

The abstract is of utmost importance for it is read by 10 to 500 times more people than hear or read the entire article. It should not be a mere recital of the subjects covered, replete with such expressions as "is discussed" and "is described." It should be condensation and concentration of the essential qualities of the paper.

Choice 2

The behavior of editor is discussed. What should be covered by an abstract is considered. The importance of the abstract is described. Dictionary definitions of "abstract" are quoted. At the conclusion a revised abstract is presented

List of References

Indexing:

TextAnalysis Copyright© MicroSystems, 1997-2002. ©2002 Megaputer Intelligence. All Rights Reserved.
The use of automated indexing tools e.g. HTML-Indexer, WebGlimpse and metadata creation tools. File://A:\Continuing Professional Development.htm
Software Tools for Indexing. webmaster@asindexing.org
Howe, Karl. Indexing. Cambridge University Press .
2004 Summer FLUG – ModernTech Electronic Delivery/ Indexing Workshop Proceeding

Abstracting:

Gopen, G.D. and Swan, J.A., 1990. The Science of Scientific Writing: Am. Scientist, 78, 550-558.
Landes, K. K., 1966, A Scrutiny of the Abstract, II: AAPG Bull., 50, 1992.
Houghton, B., 1975, Scientific periodicals: Their historical development, characteristics, and control: Linnet Books.
Urdang, L., 1973, Ed., The Random House College Dictionary, Random House, Inc.
Small, Kenneth A., "How to Write An Abstract," American Economic Association Committee on the Status of Women in the Economics Profession Newsletter (Feb. 1997), p. 22.
Michaelson, Herbert, *How to Write & Publish Engineering Papers and Report,* Oryx Press, 1990. Chapter 6 discusses abstracts.
Cremmins, Edwards, *The Art of Abstracting 2nd Edition,* Info Resources Press, April 1996. This is an entire book about abstracting, written primarily for professional abstractors.
LEO. Writing Abstract.

INDEXING

This FAQ (Frequently Asked Questions) document has been compiled to provide basic information on indexing as a freelance career, and to list resources for further information. There is a wealth of information on indexing, including books, workshops, and training courses. This file can only provide a brief orientation and a list of resources

for study. As you read it, remember that the broad, general statements are not universal laws; there are many ways to approach indexing, and people doing indexing do not fit all the broad statements in every respect. We will be happy if this file tells a non-indexer a useful way (not the best or only way) to think about the work before deciding to try to do it. We know that many (perhaps most) will realize after a time that they don't fit the rules of thumb either. If you do decide to pursue this challenging, rewarding and important work, please seek out the many other wonderful resources available; stay in touch with the field and other indexers through ASI, never stop learning and improving your skills and understanding. We hope you find this file useful. We welcome your comments.

L. Pilar Wyman (pilarw@wymanindexing.com) and Larry Harrison.

When I tell people that I am working on an index to a book, they tend to hang their heads in sorrow. I tell them that compiling an index for a book is a lot more fun than writing a book could ever be, a relaxing jaunt from A to Z compared with a jerky stop-start trek without maps.
— Craig Brown, *Times Saturday Review*, 21 July 1990

1. What is indexing?

According to the British indexing standard (BS3700:1988), an ***index*** is a systematic arrangement of entries designed to enable users to locate information in a document. The process of creating an index is called ***indexing***, and a person who does it is called an ***indexer***. There are many types of indexes, from cumulative indexes for journals to computer database indexes. This discussion concentrates on the back-of-the-book index, found in nonfiction books.

Indexes are among those necessary but never spectacular products of hard as well as skilled work that can sometimes make the difference between a book and a good book.
index review in *Books Ireland*, February 1994

The chief purpose of an index is distillation, and in performing that task it can manage to suggest a life's incongruities with a concision that the most powerful biographical stylist will have trouble matching.
Thomas Mallon, *New York Times*, 10 March 1991

The ocean flows of online information are all streaming together, and the access tools are becoming absolutely critical. If you don't index it, it doesn't exist. It's out there but you can't find it, so it might as well not be there.
Barbara Quint, ASI San Diego Conference, 1994

2. Who does indexing?

In the United States, according to tradition, the index for a non-fiction book is the responsibility of the author. Most authors don't actually do it. While a few publishers have in-house indexers, most indexing is done by freelancers, often working from home, hired by authors, publishers or packagers. (A ***packager*** is an independent business which manages the production of a book by hiring freelancers to accomplish the various tasks involved, including copyediting, proofreading and indexing.) More often, the indexer is hired by the publisher, and the fee is deducted from the money due the author. If a packager hires the indexer directly, various payment arrangements can be made.

Indexing work is not recommended to those who lack an orderly mind and a capacity for taking pains. A good index is a minor work of art but it is also the product of clear thought and meticulous care.
Peter Farrell, *How to Make Money from Home*

3. How is indexing done?

The indexer usually receives a set of ***page proofs*** for the book (images of the actual pages as they will appear, including final page numbers), often at the same time as final proofreading is being done by someone else. The indexer reads the page proofs, making a list of ***headings*** and ***subheadings*** (terms to appear in the index) and the location of each pertinent reference. After completing the rough index the indexer edits it for structure, clarity and consistency, formats it to specifications, proofreads it and submits it to the client in hard-copy form, on disk, by modem, or by email. Since the indexer is very late in the production process, there can be unreasonable time pressure.

As to *how* to index, what goes on between the ears, that's a subject for books, courses, workshops and lifelong learning from experience.

Less time is available for the preparation of the index than for almost any other step in the bookmaking process. For obvious reasons, most indexes cannot be completed until page proofs are available. Typesetters are anxious for those few final pages of copy; printers want to get the job on the press; binders are waiting; salesmen are clamoring for finished books *surely* you can get that index done over the weekend?
Chicago Manual of Style, 14th ed.

Indexers are in effect trying to provide answers to a host of unasked questions. Indexers therefore need to work as if their audience is present. But there are two snags: first, in most cases they do not know who this audience will be; second, in most cases they do not receive any feedback as to whether their judgments have been successful. From a communicative point of view, there is probably no more isolated intellectual task than indexing. The twilight howl of the indexer might well be Is there anybody there?
David Crystal, editorial, *The Indexer*, April 1995

4. Can't a computer do the indexing?

The short answer is no. Computers can easily construct a ***concordance*** (a list of words or phrases and where they appear), but this is not an index, and is not very useful to someone looking for information. The so-called automatic indexing software programs now appearing on the market are simply not up to the task of indexing a book. Book indexing involves a little bit of manipulating words appearing in a text, which computers can do, and a lot of understanding and organizing the ideas and information in the text, which computers cannot do and will not do for many years to come. An example of the difference is that a book on protective gloves for occupational use might have a chapter discussing surgical gloves, how they get punctured and how they are tested for integrity, but might never use the word *holes*. Yet a user of the book would expect to find this word in the index and be directed to the appropriate chapter. The indexer handles dozens or hundreds of such issues in every book.

Where the text is already on computer disk, the indexing features of word processing programs can ease the handling of page numbers and sorting, but the real indexing work is still done by the human. Powerful dedicated software is also available for personal computers to aid the professional indexer in constructing, sorting, editing and formatting the index, whether from hard-copy text or computer files. Many indexers use one of the programs listed on the Indexing Software page.

Automated indexing was never intended to produce back-of-the-book indexes. As Indexicon demonstrates so well, back-of-the-book indexes cannot be automatically generated.
Nancy Mulvany and Jessica Milstead, review of **Indexicon**, *Key Words*, Sep/Oct 1994

5. What skills or education do indexers need?

Many publishers and packagers don't ask for specific degrees or credentials unless they are looking for someone with subject matter expertise for a technical book. (See question 10 for more about courses on indexing.) Skills needed to learn indexing include excellent language skills, high clerical aptitude, accuracy and attention to detail. Once you are indexing professionally, you will find that self-discipline, curiosity, tolerance of isolation and love of books are necessary to keep going. In addition to all this, of course, there are the business and marketing skills needed to succeed as a self- employed professional. Clients take their cue from you: if you behave in a professional manner, most of them treat you accordingly.

Whoever the indexer is, he or she should be intelligent, widely read, and well acquainted with publishing practices also levelheaded, patient, scrupulous in handling detail, and analytically minded. This rare bird must while being intelligent, levelheaded, patient, accurate, and analytical work at top speed to meet an almost impossible deadline.
Chicago Manual of Style, 13th ed.

I wonder whether there is any profession in which a knowledge of one's tongue is of the slightest use.

T.E. Lawrence, on winning 1st place in English Language and Literature in the Senior Oxford Local Exams, 1906

6. How do indexers get clients?

Most people start by sending letters and résumés to publishers. Find their addresses in *Literary Marketplace*, *Writer's Market* and *Books in Print*, available in your library. It may take hundreds of letters to get a first indexing job. Experienced indexers say they get most jobs through recommendations from satisfied clients and networking, although some still come from marketing efforts. Now, many people interested in indexing are pursuing the apprenticeship model, working with an established indexer to build skills and experience. The three best ways to get work? Network, network, network.

7. How much are indexers paid?

That's the wrong question. (I know, I wrote the question, but that's the way it's usually asked.) A freelance indexer is running a small business; as a businessperson, you are not *paid*, you set prices and charge for a service. You are not an employee; you are an independent contractor. This is an important distinction because of how it changes your thinking. It's also very important for tax purposes, but that's off the subject. Try questions 8 and 9.

8. How do indexers price their services?

The two most common ways of quoting book index prices are per page and per entry. Different publishers prefer different methods, and indexes for different media (databases, periodicals, etc.) also are priced differently. All the different ways of quoting prices can be reduced to a fee per hour. While experienced indexers come to know what rates per page or per entry they can afford to accept, beginning indexers would be well-advised to focus on the hourly fee when figuring their bids. This enables new indexers to decide what kinds of work are best for them, and to track improvements in skill, efficiency and income as they become experienced.

If you are starting out as a freelance indexer, you won't be able to get the same fee as an indexer with 10 years of experience. This does not mean inviting exploitation by unscrupulous clients. Remember, if you are qualified as an indexer, you are producing a professional product, and you should be fairly compensated. Set yourself a rock-bottom hourly fee for run-of-the-mill indexing, the lowest figure you should ever accept, and stick to it. Remember, no one says you have to take what the prospective client is offering. No one says the client has to pay what you charge. Both parties are free to negotiate or go elsewhere. It is your responsibility to set the fee you charge for indexing, and negotiate to get it.

Indexers need to charge for their services according to the time they expect to spend on the work. On the other hand, many clients want a predictable price since they are under

budget constraints. These clients will not pay by the hour, especially if they don't know your work. How do you quote your prices to get your hourly fee?

If the client opens the discussion by saying she wants the index done for $1,200, or for a certain amount per indexable page, a fixed bid is called for. (Pricing per page is a type of fixed bid; it can be agreed to in advance even if an exact page count is not known.) Fixed bids are good for the client but risky for the indexer. The indexer must be familiar with the book before a reasonable bid can be given, because of wide variations in words per page and complexity of material. The expected number of entries per page or in the whole index should also be specified, since this is a key factor in the time spent doing the index.

Publishers in some fields (medicine, for example) want to ensure a detailed index, so they use pricing by the entry. As long as both parties are clear on exactly what constitutes an entry and how they are counted, this has the advantage of compensating the indexer for extra time spent on complex material. Again, the expected number of entries per page should be specified.

No matter how the bid is to be figured, start with the hourly fee to make sure you are being compensated according to your set rate. First, estimate how many hours it will take you to do the index, including editing, proofing and preparing final copy. This estimate is crucial. Actually indexing a representative sample of the book is helpful here, and estimating skills should improve with experience. Then multiply by your hourly fee to get the total amount you expect for indexing the book. If the client wants a lump sum bid, you are done.

To prepare a bid or price quote for a client who uses per-page pricing, divide your total estimate by the page count. To prepare a price quote on a per-entry basis, figure the total number of entries in the book and divide this into the total estimate. In summary, use one of the following methods, where $FEE is the total fee, PAGES the indexable page count of the book, and ENTRIES the total number of entries in the book (ENTRIES is average entries per page times the number of pages):

```
Fixed price =   $FEE

Per-page rate =   $FEE / PAGES

Per-entry rate =  $FEE / ENTRIES
```

If the client has a price or rate in mind, work the numbers backward to figure the hourly fee resulting from the client's number before deciding whether you can afford to accept it.

When someone offers you an indexing assignment at $12 per hour, take note of the advice from Dr. Wellisch in the next quote; you could do almost as well at McDonald's!

An hourly indexing fee should always be at least four times the wage one can earn by flipping hamburgers at a fast-food emporium.
Dr. Hans Wellisch, *Indexing from A to Z*

Hourly rates in 1993 started at $20 to $25 per hour and went up from there.
Nancy Mulvany, *Indexing Books*

9. What kind of annual income can I expect from indexing?

Here are some important factors which affect your income from indexing or any other independent service business:

1. How you set your prices.
2. How much you want to work.
3. How skilled you are at finding enough good clients to keep busy.
4. How skilled, and fast, you are at indexing.
5. How much your business expenses are.

In short, your income depends on your motivation and your business skills as well as your indexing skills. There are indexers who treat it as a relaxing, part-time business; there are indexers who work long hours and support themselves in nice middle-class style as a result. Most probably fall in between. You have to decide what you are looking for.

You need to spend time learning how to start and run a business as well as learning to index. Books and magazines on home-based business and entrepreneurship have lots of ideas and advice applicable to freelance indexers. Seminars and workshops on business skills and sales technique can be quite useful, but be careful with your money. Talk to graduates before signing up.

Suppose we look ahead to the future, finding that after gaining some experience, you reach a speed and skill level where the combination of rates paid by clients and the time you spend doing the work results in a good hourly rate. What hourly rate might you expect? ASI regularly surveys freelance and in-house indexers to find out what salaries and fees are being paid. The latest survey is available on this website, in the Members Area. If you are an ASI member, you may view and download the survey. In case you don't have access to this information, bear with me as I discuss the mechanics of estimating your self-employment income without using a specific rate.

To estimate your annual income from indexing, multiply your hourly rate by the average number of work hours in a year. Forty hours per week times 52 weeks a year is 2,080. Wait! If you want to index full-time, you need to consider all the time your business takes besides actual indexing. Writing letters and making calls to get work, rushing to the FedEx office before they close, billing, doing your tax return, shopping for supplies, backing up your computer files, meal breaks and occasional holidays and vacations

(remember those?) are all unpaid time. Don't forget idle time between jobs; it takes several years of building a client base for most indexers to get full-time work. If you plan to put 40 hours per week into your business, then allowing for all the above within the 40 hours results in a rule of thumb of about 1,200 hours per year of actual paid indexing work.

OK, multiply your hourly fee by 1,200. That's your gross revenue. But remember, this is a business; your actual income is much less. To figure hourly income, self-employment taxes (currently 15.3%) and federal, state and local income taxes come off the top, plus the cost of your supplies, utilities, ASI membership dues and the amortized cost of your office equipment. (See Schedule C of Form 1040 for calculating business taxes, expense deductions and amortization.) A good rule of thumb is to take at least 50% off the rate.

So, multiply your hourly income (about 50% of your hourly rate) by how many hours you can work per year (1,200 while you are getting established, based on a 40-hour week). If this is not enough for you to live on, don't quit the day job yet. Most indexers start indexing as a part-time moonlighting effort, supporting themselves with another job. Once they are sure they want to do this kind of work full-time, and clients are paying well and keeping them so busy it is hard to get everything done, they can make the decision to try full-time indexing.

Once you are well-established, idle periods and time spent marketing diminish, resulting in more paid work hours. You can eventually reach 2,000 hours per year, if you are willing to work more than 40 hours per week. In addition, a very good indexer who works fast can make a higher hourly rate for a given page rate, because it takes fewer hours to do the work. *Money* magazine recently published an article on successful home-based businesses which quoted one experienced freelance indexer who says he averages $50,000 per year. Some indexers are skeptical; others say this is possible after a few years if you work hard and find the right clients.

Wealth ... is more accurately measured in what you enjoy than in what you possess.
Jean Aspen, *Arctic Son*

10. How can I learn to index?

A local college or university with a Library Science or Information Science department may offer indexing courses. Many people take the indexing correspondence courses (Basic Indexing and Applied Indexing) offered by the US Department of Agriculture. Assignments are graded by indexing professionals and a certificate of completion can be provided.

If you are considering ASI membership, take note of the excellent self-paced course offered to ASI members. Available on interactive CD-ROM, this comprehensive course includes many practical exercises and self-assessments. A certificate of completion can be obtained by taking the optional exams.

Look at the Indexing Courses and Workshops web page for more information about all these options. Before you invest money in an expensive course, check out some books on the subject to gauge your interest and aptitude.

Indexing cannot be reduced to a set of steps that can be followed! It is not a mechanical process. Indexing books is a form of writing. Like other types of writing, it is a mixture of art and craft, judgment and selection. With practice and experience, indexers develop their own style as do other writers. The best we can do as teachers of indexing is to present the rules and offer guidance.
— Nancy C. Mulvany, *Indexing Books*

Index learning turns no student pale,
Yet holds the eel of science by the tail.
— Alexander Pope, *The Dunciad*

How Information Retrieval Started

The papyrus scroll used by the ancient Greeks and Romans was not the most efficient way of storing information in a written form and of retrieving it. Yet, as Greek and Roman scholars began to write large works that were compilations of data of various sorts, they found it useful to devise various means of organizing the material to make locating certain passages easier for the reader. Here are a few examples of what they did.

Tables of contents

Pliny the Elder (died 79 A.D.) wrote a massive work called *The Natural History in 37 Books*. It was a kind of encyclopedia that comprised information on a wide range of subjects. In order to make it a bit more user friendly, the entire first book of the work is nothing more than a gigantic table of contents in which he lists, book by book, the various subjects discussed. He even appended to each list of items for each book his list of Greek and Roman authors used in compiling the information for that book. He indicates in the very end of his preface to the entire work that this practice was first employed in Latin literature by Valerius Soranus, who lived during the last part of the second century B.C. and the first part of the first century B.C. Pliny's statement that Soranus was the first in Latin literature to do this indicates that it must have already been practiced by Greek writers.

Alphabetization

One method of information organization which we take for granted nowadays, namely alphabetization, was probably first devised by Greek scholars of the third century B.C. at the library of Alexandria in Egypt in order to help them organize the growing numbers of Greek literary works. If I recall correctly, the subject of alphabetization and its use in classical antiquity was treated years ago in a little monograph by Lloyd Daly.

Hierarchies of information

There are a few other ancient works which employed arranging material under headings in order to make the writing more user friendly and easier to consult.

Valerius Maximus wrote a collection of memorable deeds and sayings ca.30 A.D. The work is divided into nine books, and each book is subdivided into chapters, and each chapter has its own heading, and all entries within that chapter contain anecdotes taken from ancient literature and history which illustrate that theme.

Marcus Julius Frontinus, a Roman senator of the late first century A.D. and early second century A.D., wrote a book of military strategems in four books. Each book concerns itself with a specific area of warfare. Each book is then subdivided into chapters that each address one specific aspect of the book's major theme. Each chapter has a heading to clue the reader, and the chapter itself consists of brief extracts taken from historical works that illustrate the practical application of the topic.

Finally, Aulus Gellius wrote a work entitled *The Attic Nights* ca. 160 A.D. in 20 books. The work is a crazy quilt assortment of items on Greek and Roman history, philosophy, grammar, rhetoric, and antiquarian material in general. Since the work was composed with no real order but as the various topics occurred to the author, each chapter of every book concerns an isolated subject, and this subject is clearly spelled out in a title heading that stands at the beginning of the chapter. A reader could therefore skim through a book and locate the subject by glancing over the titles of the chapters.

Interested in more?

A brief but good discussion of the problems of ancient scholarship posed by the use of the papyrus scroll can be found on pp. 101-116 of *Varro the Scholar*, by Jens Erik Skydsgaard, published in 1968 in the series Analecta Romana Instituti Danici.

Gary Forsythe

Institute for Advanced Study

gfgf@math.ias.edu

Indexes in history

(from Hans Wellisch's *Indexing from A to Z*, H.W. Wilson Co., 1991)

Book indexes. Members of the societies of indexers may well take pride in the fact that this sense of index is indeed the oldest among the figurative or applied senses of the word, and that this specific usage (like the word itself) goes back to ancient Rome. There, when used in relation to literary works, the term *index* was used for the little slip attached to papyrus scrolls on which the title of the work (and sometimes also the name of the author) was written so that each scroll on the shelves could be easily identified without having to pull them out for inspection. "... ut [librarioli] sumant membranulam, ex qua indices fiant, quos vos Graeci ... sillybus appelatis" (so that [the copyists] may take some bits of parchment to make title slips from them, which you Greeks call sillybus) (Cicero, Atticus, 4.41.1). From this developed the usage of *index* for the title of books: "Sunt duo libelli diverso titulo, alteri 'gladius', alteri 'pugio' index erat" (There are two books with different titles, one called "The sword", the other having the title "The dagger") (Suetonius, *Caligula*, 49.3) Those two books, by the way, were what we would call today "hit lists" of people whom Caligula wished to have assassinated shortly before that same fate befell him. At about the same time, in the first century A.D., the meaning of the word was extended from "title" to a table of contents or a list of chapters (sometimes with a brief abstract of their contents) and hence to a bibliographical list or catalog...

However, indexes in the modern sense, giving exact locations of names and subjects in a book, were not compiled in antiquity, and only very few seem to have been made before the age of printing. There are several reasons for this. First, as long as books were written in the form of scrolls, there were neither page nor leaf numbers not line counts (as we have them now for classical texts). Also, even had there been such numerical indicators, it would have been impractical to append an index giving exact references, because in order for a reader to consult the index, the scroll would have to be unrolled to the very end and then to be rolled back to the relevant page. (Whoever has had to read a book available only on microfilm, the modern successor of the papyrus scroll, will have experienced how difficult and inconvenient it is to go from the index to the text.) Second, even though popular works were written in many copies (sometimes up to several hundreds),no two of them would be exactly the same, so that an index could at best have been made to chapters or paragraphs, but not to exact pages. Yet such a division of texts was rarely done (the one we have now for classical texts is mostly the work of medieval and Renaissance scholars). Only the invention of printing around 1450 made it possible to produce identical copies of books in large numbers, so that soon afterwards the first indexes began to be compiled, especially those to books of reference, such as herbals. (pages 164-166)

Index entries were not always alphabetized by considering every letter in a word from beginning to end, as people are wont to do today. Most early indexes were arranged only

by the first letter of the first word, the rest being left in no particular order at all. Gradually, alphabetization advanced to an arrangement by the first syllable, that is, the first two or three letters, the rest of an entry still being left unordered. Only very few indexes compiled in the 16th and early 17th centuries had fully alphabetized entries, but by the 18th century full alphabetization became the rule... (p. 136)

(For more information on the subject of indexes, please see Professor Wellisch's Indexing from A to Z, which contains an account of an indexer being punished by having his ears lopped off, a history of narrative indexing, an essay on the zen of indexing, and much more. Please, if you quote from this page, CREDIT THE AUTHOR. Thanks.)

Indexes go way back beyond the 17th century. The Gerardes Herbal from the 1590s had several fascinating indexes according to Hilary Calvert. Barbara Cohen writes that the alphabetical listing in the earliest ones only went as far as the first letter of the entry... no one thought at first to index each entry in either letter-by-letter or word-by-word order. Maja-Lisa writes that Peter Heylyn's 1652 *Cosmographie in Four Bookes* includes a series of tables at the end. They are alphabetical indexes and he prefaces them with "Short Tables may not seeme proportionalble to so long a Work, expecially in an Age wherein there are so many that pretend to learning, who study more the Index then they do the Book."

Other articles on the history of indexing:

"The Oldest Printed Indexes," Hans Wellisch, *The Indexer* 15 (2), pp. 73–82, 1986.

"Early Humanist Indexing," Elizabeth Eisenstein, *The Printing Press as an Agent of Change* quoted in *The Indexer* 14 (1), p. 58, April, 1984.

"Indexing," Hans Wellisch, in E*ncyclopedia of Library History*, Wiegand and Davis, pp. 268–270.

INDEXES and INDEXERS in FICTION

Ballard, J. G. *War fever.* London : Collins, 1990. ISBN: 0002237709.
The index itself could be considered fiction. See information about this index in The Indexer, volume 17, issue 4, page 250.
Balzac, Honoré de *La comédie humaine*
See information about this index in The Indexer, volume 16, issue 4, page 246.

Bradbury, Malcolm *My strange quest for Mensonge : structuralisms hidden hero.* London : A. Deutsch, 1987. ISBN: 0233980202.

See information about this index in The Indexer, volume 16, issue 3, page 194. and volume 16, issue 4, page 248.

Card, Orson Scott "The Originist"

It was originally published in *Foundation's Friends* (ed. by Martin Harry Greenberg, 1989. ISBN 0312931743). It is also in *Maps in a Mirror* (Orson Scott Card, 1990. ISBN 0312850476) a huge hardcopy compilation of Card's works, and in *Flux* (a paperback version of one section of *Maps in a Mirror*).

In "The Originist", the main character's wife is an indexer in the Foundation library, and the fact that she is an indexer does not come into play until near the end of the story, but her work turns out to be crucial to the research her husband is doing. The story is excellent reading, and the environment in which the indexer works is very inspiring.

Card, Orson Scott "Homecoming" trilogy of novels: *The Memory of Earth* (1992, ISBN 0312852797 or 0312930364), *The Call of Earth* (1993, ISBN 009926191X or 0099260115), and *The Ships of Earth* (1994, ISBN 0312856598).

One of the key plot elements involves the theft and use of the Index to the Oversoul. The Oversoul is the omniscient computer which has guided human development on the planet Harmony for millions of years. Parts of the Index [i.e., directory] to the Oversoul are lost, and the Oversoul realizes that significant parts of its memory have been lost through the ages - inaccessible without the Index.

Carroll, Lewis *Sylvie and Bruno*

See information about this index in The Indexer, volume 7, issue 2, page 55 and volume 18, issue 2, page 110 and issue 20, volume 1, pages 11-13.

Crewe, Candida *Mad about Bees* Heinemann, 1991.

Defoe, Daniel *Memoirs of a cavalier*. London, Oxford University Press, 1972. ISBN: 0192553611 and 0192553631.

See information about this index in The Indexer, volume 16, issue 4, page 245.

Doyle, Arthur Conan, Sir *Sherlock Holmes and the Creeping Man*, *The Adventure of the Sussex Vampire*, *A Scandal in Bohemia*, *A Case of Identity*, and *The Case of the Empty House*.

In these stories, Sherlock Holmes makes use of a thorough personal index. Take a look...

"Make a long arm, Watson, and see what V has to say." I leaned back and took down the great index volume to which he referred. Holmes balanced it on his knee, and his eyes moved slowly and lovingly over the record of old cases, mixed with the accumulated information of a lifetime. "Voyage of the Gloria Scott," he read. "That was a bad business... Victor Lynch, the forger. Venomous lizard or gila. Remarkable case, that! Vittoria, the circus belle. Vanderbilt and the Yeggman. Vipers. Vigor, the Hammersmith wonder. Hullo! Hullo! Good old index. You can't beat it. Listen to this, Watson.

Vampirism in Hungary. And again, Vampires in Transylvania." Excerpt from *The Adventure of the Sussex Vampire*

Ellmann, Lucy *Sweet Desserts.* Viking Penguin, 1989. ISBN 0670827118.

Another example of an author-written "index". This one doesn't really index the book -- rather, it is part of the text, expanding the characters and the humour. It's a story of two sisters, written by one. Some examples of entries: **Boyfriends, Franny's hand-me-down. *see* less of** and **Cat, unforeseeable responsibilities relating to ownership of a** and **Divorce, cut your losses and go to Las Vegas**. See information about this index in The Indexer, volume 16, issue 3, page 194 and volume 16, issue 4, page 248.

Franken, Al *Rush Limbaugh Is A Big Fat Idiot and other observations.* New York : Delacorte Press, 1996.

Franken wrote this entire index himself. It is totally inaccurate and hysterical. Sample entries? **Dirigible, Limbaugh size of** and **Doughnuts, Limbaugh's consumption of**.

Gaarder, Jostein *Sophie's world*. London : Phoenix, 1995 and 1991. ISBN: 1857993284.

This is a compilation of lectures on the history of philosophy linked by a narrative about a young girl, and marketed as a novel. It has an index that deals only with the philosophical content, not to the fictional -- and no note explaining this discrimination.

Graves, Michael A. R. *Thomas Norton : the Parliament man.* Oxford, UK ; Cambridge, Mass. : Blackwell Publishers, 1994. ISBN 0631167994.

This biography of Norton, a 16th Century English author and legislator, does mention his involvement in indexing, classifying, and cataloging.

James, Clive *Brilliant creatures*. London : J. Cape, 1983. ISBN: 0224021222.

See information about this index in The Indexer, volume 13, issue 4, page 277 and volume 16, issue 4, pages 244, 246.

James, P.D. *Death of an Expert Witness.* New York : Scribner, 1977. ISBN 0684152673.

The villain chooses the blackmail victims by postcoordination of their psychological symptoms as indexed on edge-notched cards.

Kaminsky, Stuart *Blood and Rubles* Wheeler Publishing, 1996. ISBN 1568953291 and 0804112886.

Emil Karpov, the character with no emotions, is looking for one Igor Kuzen. Karpov has kept notebooks detailing every unfinished case he ever worked on in hopes of finally nailing the ones who got away. He has bookshelves lined with these notebooks.
"...He continued his search. There were Igor Kuzens listed in the directory, and the MVD computer system had come up with a probable Igor Kuzen, a medicine hijacker, but he was in prison. The name had touched a memory in Karpov. He had seen it somewhere, written it somewhere, and now he was methodically going through his cross-index in search of a reference. All names listed in his books of notes were cross-indexed. He couldn't find it.
... "Spelling," he said aloud, flipping through the index volume where each entry was clearly printed in his own precise hand. He was now going through the **Ts**, and that was where he found it. Igor Tuzen. A single reference. July 1986. Questioned in relation to the beating and death of a woman who lived in the apartment next to his. The man had identified himself as a physicist..."

Nabokov, Vladimir *Pale fire*. New York : Putnam, 1962.

This book is not about an indexer, but it does include a "joke" index. See information about this index in The Indexer, volume 7, issue 2, page 55 and volume 12, issue 4, page 200.

Ormesson, Jean d' *La gloire de l'Empire*. Paris : Gallimard 1971.

See information about this index in The Indexer, volume 8, issue 2, page 118.

Proust, Marcel *A la Recherche du Temps Perdu*. [S.l.] : Gallimard, 1954.

See information about this index in The Indexer, volume 7, issue 3, page 133 and volume 10, issue 3, page 159 and volume 16, issue 4, page 245-246.

Pym, Barbara *No Fond Return of Love.* Portway, Bath, [Eng.] : C. Chivers, 1961. ISBN 0855949309.

Many of Pym's works include indexers.

Renault, Mary *The lion in the gateway*. 1964

See information about this index in The Indexer, volume 16, issue 4, page 246.

Scott, Walter, Sir *Waverley novels*. Portrait Edition (1913) and A&C Black editions (1886-7).

See information about this index in The Indexer, volume 8, issue 3, pages 153-9 and volume 10, issue 4, page 177 and volume 16, issue 4, page 245.

Tolkien, J. R. R. *The Lord of the Rings.* ISBN 0395193958.

The index in this text has a long and somewhat elusive history. Here's what scholars know...
In a biography by Humphrey Carpenter (originally published: London : Allen & Unwin, 1977. ISBN 0261102451 (pbk).), we read that for the first edition "He had abandoned any hope of making an index of names, having found that the job would take too long." Carpenter's book of Tolkien's letters (originally published: London : Allen & Unwin, 1981. ISBN 0044406649 (pbk).) mentions a revision for a second edition that was to thwart the pirated paperback edition in the U.S. "I am hoping that alteration of the introductions, considerable modifications of the appendices and the inclusion of an index may prove sufficient for the purpose." From the introduction in the revised edition of this text (*The Fellowship of the Ring*, Houghton Mifflin, 1982. ISBN 0395193958), we see this note: "this edition offers this Foreword, an addition to the Prologue, some notes, and an index of the names of persons and places. This index is in intention complete in names but not in references, since for the present purpose it has been necessary to reduce its bulk. A complete index, making full use of the material prepared for me by Mrs. N. Smith, belongs rather to the accessory volume." The "accessory volume" he refers to did not at that point exist.

For more information about this index in The Indexer, volume 7, issue 2, page 55.

Trollope, Anthony "The Spotted Dog" in *An Editor's Tales.*

First published in St Paul's Magazine, then in book form in 1870; reissued by Penguin, 1993 (pages 235-328).

Vonnegut, Kurt *Cat's Cradle.* New York : Delacorte Press, 1963. ISBN 0385281269.

In this book, the indexer makes a comment that you can tell a lot about a person by how they index a book.

Vonnegut, Kurt *Jailbird*. Cape edition (1979), and Granada edition (1981).

See information about this index in The Indexer, volume 12, issue 2, page 109 and volume 12, issue 4, page 222 and volume 16, issue 4, page 246.

THESAURUS INFORMATION

WHAT IS A THESAURUS?

- **For writers**, it is a tool like Roget's — one with words grouped and classified to help select the right word to convey a specific nuance of meaning.
- **For indexers and searchers**, it is an information storage and retrieval tool: a listing of words and phrases authorized for use in an indexing system, together

with relationships, variants and synonyms, and aids to navigation through the thesaurus.

For more information on what a thesaurus is and what it contains, see: National Information Standards Institute. *American National Standard Guidelines for the Construction, Format, and Management of Monolingual Thesauri*. Bethesda, MD: NISO Press, 1994. (ANSI/NISO Z39.19-1993).

WHEN DO I NEED A THESAURUS?

- **For a single back of the book index**, the terms derived from the text, plus the cross-references that are added, represent the "thesaurus" for that index, so developing a separate thesaurus probably isn't worthwhile. However, you may find an existing thesaurus helpful for terminology questions.
- **For a large index**, one that is ongoing, or one where multiple indexers will be involved, a thesaurus is almost a necessity. There is no way to maintain consistency over time or across indexers without one.

CAN I USE AN EXISTING THESAURUS?

- Almost all existing published thesauri have been developed to serve the needs of a particular database. Unless your index is of exactly the same scope and level, an existing thesaurus is unlikely to meet your needs precisely.
- However, there may be an existing thesaurus that you can adapt or even use as is, and it's worth looking.

Different fields vary in how well they are served by thesauri. With the exception of the *Art and Architecture Thesaurus*, science and technology are far better covered than the humanities, with the social sciences falling somewhere in between.

And if the problem weren't complicated enough, just finding out if a thesaurus exists for a subject area can be a major challenge. *The Thesaurus Guide*, 2nd ed., 1992, published by the Commission of the European Communities, is immense, but very incomplete. Another possibility is to check out databases in the subject area, and find out if they have published thesauri.

The above information was written by Jessica Milstead specifically for use on the ASI Web site.

HOW DO I BUILD A THESAURUS?

THE TOP-DOWN METHOD:

- Convene a group of subject experts to decide on the scope and broad categories of terms to be included.

- Use existing dictionaries and thesauri to decide on the terms and their relationships.
- Review and organize the preliminary term set: decide on preferred terms and make Use references from the variants and synonyms; and build hierarchical and associative relationships among the preferred terms.
- Produce a draft thesaurus, test index and revise.

THE BOTTOM-UP METHOD:

- Develop a group of subject experts to serve as advisors; work with them to determine the scope if it is not already set.
- If there is a set of representative already-indexed documents, use the index terms from this set as your preliminary term list.
- If not, index a set of representative documents using free language (i.e., no vocabulary control), and take this term set as your preliminary list.
- Build your thesaurus by reviewing and organizing these terms, using a variety of resources as aids, as in the top-down method.
- Refer to your subject experts on terms whose meaning or usage is unclear, and for advice on which variant or synonym to prefer (or on whether two terms really are synonyms in the field).
- Produce a draft thesaurus, test index, and revise.

MAINTAINING YOUR THESAURUS

A thesaurus is never "finished," unless it is no longer being used for indexing or its database is no longer being updated. Plan for maintenance before you even begin developing your thesaurus. A thesaurus which is not well-maintained quickly becomes a liability rather than an asset.

FOR MORE INFORMATION...

- *American National Standard Guidelines for the Construction, Format, and Management of Monolingual Thesauri*. Bethesda, MD: NISO Press, 1994. (ANSI/NISO Z39.19-1993, ISSN 10415653, ISBN: 1880124041).
- Aitchison, Jean, David Bawden, and Alan Gilchrist. *Thesaurus Construction and Use: A Practical Manual*. 3rd ed. Chicago/London: Fitzroy Dearborn, 2000. (ISBN 1579582737).
- Craven, T. Introductory tutorial on thesaurus construction. London, Ontario, Canada: University of Western Ontario, Graduate School of Library and Information Science. Step by step guidance on constructing a thesaurus, including collecting terms, deciding on the forms of terms, and choosing the appropriate relationships with which to link terms.

Thesaurus Management Software

If you're thinking about using database software you already have, or even building a thesaurus using word processing software, DON'T. Even a small thesaurus will represent a very large investment of time and intellectual effort. Software which will automate the clerical and repetitive tasks is available for a cost that's very reasonable when you consider the true cost and value of the tool you will be producing.

A listing of some of the thesaurus management packages which are on the market today follows. The listing is organized in two parts: standalone software, and modules which are integral parts of larger systems.
Don't be confused by statements about "has a thesaurus" in the publicity for packages designed for other purposes such as retrieval or library automation.

Many of these don't manage a thesaurus; they simply provide access to a thesaurus file that's stored in the package or read in from another resource.

Thesaurus Management Software Packages

This list of software geared toward the needs of professional indexers and vocabulary developers is for informational purposes only. It is not intended to be a comprehensive list of such tools. As always, neither ASI nor the author endorses or promotes products, companies, political parties, religious groups, or other sites and information provided therein.

Standalone packages

Except as noted, these packages run on PCs and/or are available via the Web. Some also run under Unix; check with the vendors for specifics.

a.k.a.® Classification Software

Developed in accordance with ISO 15489 standards, **a.k.a.®** is Australia's foremost thesaurus software and taxonomy builder.

a.k.a.® Classification Software was launched in August 2000 and the response from the market has been overwhelmingly enthusiastic. **a.k.a.®** is now used by government departments, city councils, utilities, legal firms, universities, hospitals and financial institutions and has become the product of choice for many major developers of keyword thesauri and classification schemes.

a.k.a.® evolved out of Synercon's need for an industrial strength thesaurus builder for their consultants and clients to use. Says **Conni Christensen**, Principal Consultant: We searched the market but couldn’t find a thesaurus tool that really met our own needs. We wanted something that was simple to use, but also published classification schemes to a range of records and document management

systems, including the Windows folder structure. As we couldn't find an existing thesaurus tool to match our specifications, we decided to develop our own.

a.k.a.® Classification Software delivers:

- ISO 15489 Business Activity Classification Schemes
- Keyword in context publication
- Thesaurus builder
- Glossary builder
- Case file term management
- Controlled lists linked to case file terms
- Integrity reports to identify orphans and duplicates
- Global spell checking
- Integration with **a.k.a.® Disposition Management Software**

See the complete list of **a.k.a.®** Classification Software functionality.

Data Harmony

Offering abstracting, indexing, and thesaurus creation services, Access Innovations has served the business, government, academic, library, and research communities with innovative information solutions since 1978.

Since 1998, we have been providing these communities with the same tools that we use in-house. Data Harmony software is production line hardened; we use it daily in the creation of thesauri and tagging of information sources for our customers. You can use this software yourself to make and maintain taxonomies and indexing systems that lead users to success in their quest for information.

Unleash the Power of Your Taxonomy

- Create and manage controlled vocabularies
- Transform your taxonomy into a full thesaurus or ontology
- Craft your thesaurus to support navigation and search
- Enrich your vocabulary with facets, custom relationships, and RDF tags
- Define attributes to suit your requirements
- Import and merge vocabularies
- Export in XML, SKOS, OWL, and 11 other formats
- Map your thesaurus to another thesaurus

- Make your thesaurus multilingual
- Tailor your systems to ensure relevant and complete search returns.
- Integrate human intelligence into automated indexing
- Automatically index legacy collections
- *Make search smarter*

The Data Harmony line of software products from Access Innovations is now available on the GSA Schedule. Follow the link below or call our Customer Service line to learn more.

GSA Advantage!
www.gsaAdvantage.gov
Access Innovations [GS-35F-0820N]

Customer Service
1-800-926-8328

Thesaurus and Indexing Tools

Offering abstracting, indexing, and thesaurus creation services, Access Innovations has served the business, government, academic, library, and research communities with innovative information solutions since 1978.

Since 1998, we have been providing these communities with the same tools that we use in-house. Data Harmony software is production line hardened; we use it daily in the creation of thesauri and tagging of information sources for our customers. You can use this software yourself to make and maintain taxonomies and indexing systems that lead users to success in their quest for information.

Access Innovations Inc.®

MultiTes

MultiTes

Quick info... MultiTes Pro MultiTes WDK Managed Web Solutions Consulting **See what clients are saying...** All... Australia Belgium Brazil Canada Costa Rica Germany The Netherlands Philippines United Kingdom United States

Thesaurus Construction and Publishing Solutions

Since 1983, our tools have made it easy to create and manage thesauri, taxonomies and other types of controlled vocabularies. Our products combine more than 20 years of thesaurus software experience with the latest advances in web technology. As more innovations become available, we have expanded our services to include managed publishing of thesauri on the Internet and custom solutions for corporate needs.

Software tools to build and publish thesauri

All our products are based on the same core thesaurus technology which has been tested and proven reliable for many years. As new technology and new standards become available, we enhance our products to take advantage of new features and help thesaurus builders stay up to date with trends.

Our software products include:

- MultiTes Pro (Workstation)
- MultiTes WDK (Publish thesauri on the Internet/Intranet)
- MultiTes EDK (Deploy thesauri on your corporate servers or Internet application)

Managed web solutions for your thesauri

Our managed web solutions provide you with the mechanisms to publish thesauri on the web, without having to worry about network administration, firewall settings, server management, hacker intrusions, etc.. Just send us your thesaurus and we will set it up on the web.

We offer two types of services:

1. Thesaurus publishing
2. Thesaurus construction and management

STRIDE

STRIDE is one of the most powerful and versatile thesaurus management systems available. It builds and maintains standard vocabularies for use in information indexing and retrieval. It is simple to use and gives rapid response even when accessing very large thesauri.

- STRIDE supports all the thesaurus relationships and display fornats defined in the standards ISO 2788 and ANSI Z39.19 and enables you to define your own special relationships as well.
- By setting up suitable relationships you can implement full multi-lingual thesauri, as specified in ISO 5964.
- STRIDE gives you full control over how much information STRIDE displays, using a simple but flexible display language. In addition to the standard displays, STRIDE provides symbolic and graphical tree representations of term relationships. You have full choice of font and colour for your display and all displays can be printed. .
- You can easily create or remove one-to-many links using dialogs and buttons. STRIDE automatically creates the reciprocal links and maintains an integrity check on all your linking.
- STRIDE has hypertext facilities which enable you to move to related contexts while scanning a thesaurus.
- STRIDE is convenient for the single PC user but it also has all the features required for a major multi-user system. It allows both interactive and batch editing and provides facilities for automatic logging of transactions. Thesaurus modifications take immediate effect, ensuring instant currency of the updated thesaurus for all users.

Synaptica

Synaptica from Dow Jones simplifies and standardizes business vocabulary management. Used by clients in the Corporate, Pharmaceutical, Online Publishing, Government, eCommerce and Internet Search Markets, Synaptica enables organizations to add structure and value to existing information assets and connects professionals to the information they need, when and where they need it.

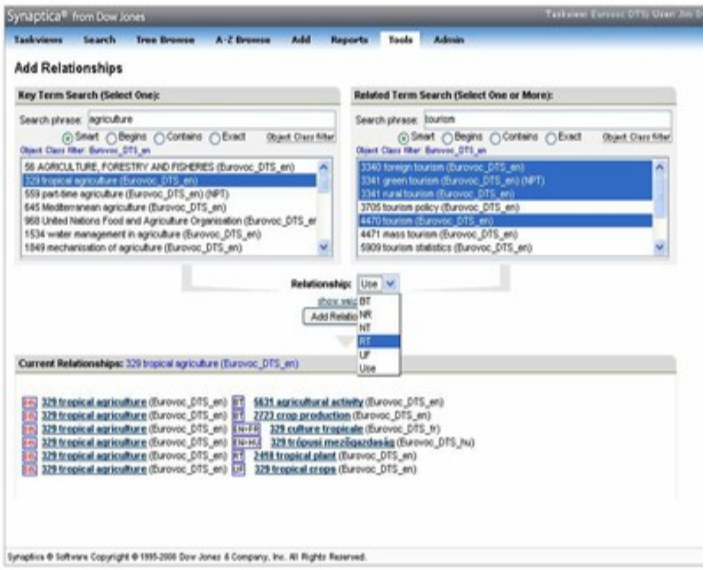
Synaptica is used to efficiently build and maintain enterprise taxonomies, thesauri, name catalogs and other authority files. With Synaptica, you can design new metadata and vocabulary screens, data elements, and rules - in minutes and easily manage hundreds of vocabularies, millions of terms and unlimited users with granular permission controls.

Synaptica is the result of over a decade of research and development driven by both power users in the information community and the evolution of the Semantic Web. An unparalleled set of features has

evolved to meet the uncompromising demands of global taxonomy and authority file experts.

With easy configuration, audience-centric flexibility and valuable governance tools, Synaptica has solved information management challenges for many global Fortune 500 organizations. Synaptica enables you to:

- Drive online revenue and increase customer satisfaction by improving how online customers find your information, products or services.
- Improve product development with a global business vocabulary that feeds right into downstream applications such as portals, reporting programs and CRMs.
- Standardize your corporate vocabulary and improve information management to streamline product development, regulatory compliance and risk mitigation.
- Monetize media content by unifying digital platforms to enable a better search infrastructure and a competitive advantage in the media marketplace.
- Gain valuable productivity benefits by connecting users and workgroups with the relevant information they need, within corporate intranets or portals.



Term relationships are now displayed side-by-side, making it easy for the user to find and edit a key term and multiple related terms. Just select your relationship class and from there you may edit any existing or new terms and relationships.

Term Tree 2000

World ranking thesaurus and classification software

Press the product name for more information or the associated Try Now button to request a demonstration kit.

Term Tree  One-2-One

Term Tree, world ranking thesaurus software, featuring a Windows Explorer interface, to continually show all relationships. Term Tree includes features to create and manage subject *(ISO 2788)* or records management scheme *(AS4390 - ISO 15489)* thesauri, synonym rings, controlled vocabularies or taxonomies.

More information is available on the Term Tree product page.

One-2-One. The next generation classification management software designed with **advanced features for records management** EDRMS requirements while still retaining the flexibility of a standalone development tool.
Develop, then maintain, a new hierarchical classification, file-plan, thesaurus or classification in standalone mode or, optionally, via direct connect (using API/SDK) with supported target product. Currently supporting Trim Context and Livelink eDocs [previously Hummingbird Enterprise (2004 or 2005)]. No need to export to transfer files then import and verify.

Click link to view
sample exported thesaurus web pages or sample exported hierarchical web pages

Database modules

In general, these modules are integral parts of the larger system, and cannot be run separately. Their availability may vary, depending on the vendor's development priorities.

TheMa (Topic) - thesaurus manager for Oracle

TheMa is a professional, powerful, web based management tool for Oracle® thesauri.

Full text search today is a standard in most archiving systems. Unfortunately, only a small part of the power of modern database systems is utilized in many applications.

Oracle® database systems offer functionality far beyond simple "free text search". For example they provide means for stemming or thesaurus based queries. Thesaurus based queries for example enables you to find documents containing terms which are in one way or another related or similar to your search term. This feature allows for much more sophisticated work with your data.

In order to build up and continuosly manage a thesaurus, you need an efficient tool which you can learn use easily. Our Thesaurus Manager for Oracle® , called ***TheMa***, provides an easily used, yet powerful administration interface to your Oracle®-based thesaurus. It supports large thesauri and foreign language character sets.

New: TheMaView

As a useful add-on to ***TheMa,*** TRIGA has introduced a module called TheMaView. It enables you to utilize your thesaurus within an existing web site without major changes to it's software. Details are provided here.

New: Version 1.5

Starting February 2008, we deliver the new version ***TheMa 1.5***. Registered customers find the new package in our download area. An overview of this version is available here.

THESAURI ONLINE

There are many thesauri available through the Internet. Below is a list of those that visitors to our site have found most useful. If you would like to suggest an addition to this page, please contact webmaster@asindexing.org.
Aquatic Sciences and Fisheries Thesaurus (browse only)
A link to this thesaurus is found at the Cambridge Scientific Abstracts web site.Astronomy Thesaurus (browse and search)
This thesaurus, compiled by Robyn M. Shobbrook & Robert R. Shobbrook provides access to the thesaurus in English, French, German, Italian, and Spanish.
The Art & Architecture Thesaurus (browse and search)
The AAT is a controlled vocabulary for describing and retrieving information on fine art, architecture, decorative art, and material culture.
Environmental Data Catalogue Thesaurus (search only -- German and English) The International Board of the 'Wortgutredaktion' of the

"Thesaurus of the Environmental Data Catalogue" (WGR) was initiated to continually adapt the contents of the Thesaurus of the Environmental Data Catalogue to the needs of its users. The Board is composed by representatives of institutions, environmental experts and co-operation partners, and it is directed by the "Co-ordination Unit Thesaurus Development" at the Federal Environment Agency of Austria in Vienna.

ERIC (Educational Resources Information Center) Thesaurus (search only) ERIC is a national education database of bibliographic citations and digests, sponsored by the U.S. Department of Education. This online version of the ERIC Thesaurus provides immediate search access to portions of the ERIC database. The Thesaurus content is the result of 30 years of work and refinement by the ERIC system, a decentralized network of clearinghouses and support components. The 40-member ERIC Vocabulary Review Group (VRG) presently oversees the Thesaurus development effort.

The Getty Thesaurus of Geographic Names (browse and search)
The TGN contains approximately 900,000 records for places, arranged in hierarchies representing all nations of the modern world, and including vernacular and historical names, coordinates, place types, and other relevant information.

International Thesaurus of Refugee Terminology (ITRT)
From the United Nations High Commissioner for Refugees Library (UNHCR) and Forced Migration Online (FMO). It is as an interactive and searchable tool online for the location of relevant indexing terminology in English, French or Spanish.

Kinsey Institute
The Kinsey Institute has at its site a thesaurus that deals with all aspects of sexuality. Although there is no direct link to their thesaurus, you can contact the librarian of the site for more information.

Library of Congress Thesauri
Links to a variety of thesauri, including the Legislative Indexing Vocabulary (LIV).

LC Thesaurus for Graphic Material (browse and search)
If you are looking for a thesaurus of terms related to graphic materials, this is one you'll want to explore.

Life Sciences Thesaurus (browse only)
This Fourth Edition of the Life Sciences Thesaurus contains some 9,400 descriptors used in indexing documents for the Life Sciences databases and corresponding printed abstracts journals.

Medical Subject Headings (MeSH) (browse only)
This thesaurus is brought to you by the National Library of Medicine. MeSH is the most used thesaurus in publications for medical and health professionals.

NASA Thesaurus (browse only)

Although this is the NASA Thesaurus, it includes much more than just aerospace terms (for example, it includes "Abdomen" with broader terms and related terms).
National Monuments Record Thesauri (browse only)
Standardisation of terms to be used when creating new records of the past. Seven separate online thesauri are available.
Thesaurus of Sociological Indexing Terms (browse only)
UNBIS Thesaurus (browse only)
This online database represents the fourth edition of the UNBIS Thesaurus and the first in all the official languages of the United Nations: Arabic, Chinese, English, French, Russian, and Spanish. The multilingual UNBIS Thesaurus, created by the Dag Hammarskjöld Library, United Nations Department of Public Information, contains the terminology used in subject analysis of documents and other materials relevant to United Nations programmes and activities. It is used as the subject authority of the United Nations Bibliographic Information System (UNBIS) and has been incorporated as the subject lexicon of the United Nations Official Document System. It is multidisciplinary in scope, reflecting the Organization's wide-ranging concerns.
UNESCO Thesaurus (browse only)
The UNESCO Thesaurus is a controlled vocabulary developed by the United Nations Educational, Scientific and Cultural Organisation which includes subject terms for the following areas of knowledge: education; science; culture; social and human sciences; information and communication; and politics, law and economics. It also includes the names of countries and groupings of countries (political, economic, geographic, ethnic and religious, and linguistic groupings).
The Union List of Artist Names.
The ULAN is a database of biographical and bibliographical information on artists and architects, including a wealth of variant names, pseudonyms, and language variants.

Indexing Evaluation Checklist

The Index is the KEY to the book

Is the index to your book or web site good enough for your readers?
Here are some helpful insights for ensuring an excellent index.

"An index is not an outline, nor is it a concordance. It's an intelligently compiled list of topics covered in the work, prepared with the reader's needs in mind."

Reader Appropriateness

- Are the indexed terms appropriate for the intended audience? For example: "heart attack" in a book for the general public, "myocardial infarction" in a book for health professionals; "Taxus" in a work for botanists or horticulturalists, "Yew" in a work for home gardeners.

Main Headings

- Are the main headings relevant to the needs of the reader? Are they pertinent, specific, comprehensive? Not too general yet not too narrow? Not inane or improbable?

- Do main headings have not more than 5–7 locators (page references)? If more, they should be broken down into subheadings.

Subheadings

- Are the subheadings useful? In the example below,
  a) the page ranges are extensive
  b) the subheading "problems with Republicans" may be too general Roosevelt, Franklin problems with Republicans, 1–32
- Are subheadings concise, with the most important word at the beginning? For example, not: banks and relationship to Federal Reserve bank but banks
  Federal Reserve regulation
- Unnecessary words and phrases like "concerning" and "relating to" and proliferation of prepositions and articles should be avoided.

- Is the number of subheadings about right? More than one column's worth is probably too many. Are subheadings overanalyzed? Could they be combined? For example, could "dimensions" be substituted for "height," "width," and "length"? Or should some subheadings become main headings with their own subheadings instead?

- Do subheadings have more than 5–7 locators? If more, they should either be broken down into sub-subheadings or be changed to main headings.

Double Postings

- For the reader's convenience, many subheadings should be double posted—that is, they should exist as main headings too. An example: "Cats: Siamese" and "Siamese cats." Has this been done? Double postings should, of course, have the same locators. Do they?

Locators (Page References)

- Are the locators accurate? Check a sample of entries to see. Spot-check pagination for nonsense numbers where the hyphen or en dash may be missing, such as 18693 for 186-93. Check that elision (page ranges such as 186-93) is consistent.

- When locators include roman numerals or volume numbers, does the typography make the usage clear?

Cross-References

- Have ***see*** and ***see also*** cross-references been provided?
- A ***see*** should direct the reader to a different term expressing the same concept, such as "Clemens, Samuel. ***See*** Twain, Mark" or "aerobics ***see*** exercise".

- A ***see also*** should guide the reader from a complete entry to the related entries for more and different information. Examples: "Mammals: 81, 85, 105; ***see also*** names of individual mammals" "astronomy 12–14, 56, 68. ***See also*** galaxies; planets"

Length and Type

- Is the index length adequate for the complexity of the book? An index should be 3–5% of the pages in the typical nonfiction book, perhaps 5–8% for a history or biography, and more (15–20%) for reference books.

- Is there a need for more than one type of index? For example, in addition to the usual subject index, perhaps a separate name or place index is called for. If so, is there one?

Format

- Is the type large enough to be easily read? Do the index pages look open and not crowded?
- Are the main headings and subheadings (and sub-subheadings if any) distinguished from each other?
- Is the organization—whether alphabetical, chronological, or other—accurate, clear, and consistent?
- When an entry's subheadings "turn a page" that is, are continued from a right-hand page to a left-hand page, the main heading should be repeated, followed by the word *continued* in parentheses. Depending on the size of the pages, continued headings might be appropriate for continuations from left to right pages, or even from left to right columns. Are they present?

- Preferences for punctuation between main headings and their subheadings and ***see*** and ***see also*** cross-references will vary from publisher to publisher. This discussion features several acceptable variants. The important thing is that the punctuation style be clear to the reader and consistent. Is it?

The National Library of Medicine (NLM) Medical Subject Heading (MeSH) Indexing System

What is MeSH?

- Acronym for **Medical Subject Headings**
- The U.S. National Library of Medicine's controlled vocabulary (thesaurus)
- Gives uniformity and consistency to the indexing and cataloging of biomedical literature and is a distinctive feature of MEDLINE
- Similar to keywords on other systems
- Arranged in a hierarchical manner called the MeSH Tree Structures

- Updated annually

Who uses MeSH?

- Searchers of MEDLINE/PubMed, library catalogs, and other databases -- to assist with subject searching
- National Library of Medicine (NLM) indexers -- to describe subject content of journal articles for MEDLINE
- Catalogers -- to describe books and audiovisuals in the NLM and other medical library collections

MeSH Vocabulary includes four types of terms:

- Headings — over 24,000 headings represent concepts found in the biomedical literature

  Examples:
  - Body Weight
  - Kidney
  - Dental Cavity Preparation
  - Self Medication
  - Radioactive Waste
  - Brain Edema

- Subheadings — (also called qualifiers) attached to MeSH headings to describe a specific aspect of a concept

  Examples:
  - adverse effects
  - diagnosis
  - metabolism
  - therapy

  Note: See Subheading Hierarchical Groupings

- Supplementary Concept Records — over 172,000 terms in a separate chemical thesaurus [updated weekly]

  Examples:
  - cordycepin
  - valspodar
  - tacrolimus binding protein 4

- Publication Characteristics (Publication Types) — describe the type of publication being indexed; i.e., what the item is, not what the article is about

  Examples:
  - Letter
  - Review
  - Randomized Controlled Trial

The MeSH® Tree Structures

MeSH headings are organized in a "tree" with 16 main branches:

A. Anatomy
B. Organisms
C. Diseases
D. Chemical and Drugs
E. Analytical, Diagnostic and Therapeutic Techniques and Equipment
F. Psychiatry and Psychology
G. Biological Sciences
H. Physical Sciences
I. Anthropology, Education, Sociology and Social Phenomena
J. Technology and Food and Beverages
K. Humanities
L. Information Science
M. Persons
N. Health Care
V. Publication Characteristics
Z. Geographic Locations

Each branch has many levels of sub-branches, and each heading has a position in the hierarchy.

Anatomy
Body Regions
Back
Lumbosacral Region
Sacrococcygeal Region

Some terms appear in more than one branch of the tree.

Anatomy	Anatomy
Body Regions	Sense Organs
Head	**Ear**
Ear	Ear, External +
	Ear, Inner +
	Ear, Middle +

The hierarchy allows a MEDLINE/PubMed search of a broader term to include the narrower terms in all branches automatically. This is known as “exploding.”

Subheading Groupings

- Subheadings are arranged in logical hierarchical groupings (families).

- Not all subheadings are placed in these groupings as they do not logically fit.

Principles of MEDLINE Subject Indexing

To better understand how to use MeSH when searching MEDLINE, it is helpful to understand the basic principles of subject indexing for MEDLINE.

Subject indexing includes:

- reviewing a journal article (or other material such as a letter or editorial)
- determining its subject content, and
- describing that content using a controlled vocabulary.

Purpose of indexing with controlled vocabulary:

- to facilitate search retrieval by eliminating (or accounting for) the use of variant terminology for the same concept.

Types of vocabulary terms:

- MeSH headings ("descriptors")
- Subheadings ("qualifiers")
- Check tags
    - A special class of MeSH headings that must be considered routinely for every article
        - Primarily headings covering species (including Humans), gender, and human age groups
        - Also cover:
            - historical time periods
            - pregnancy
    - Usually are indexed even if merely mentioned, unlike other MeSH headings
    - For review articles, usually are indexed only if the main point of the article
    - Humans is not selected for articles involving institutions that serve humans, e.g., clinics, hospitals
    - Some check tags (Animals, Humans, Male, Female, In Vitro, Comparative Study, and research support terms) can never be designated as the main point of the article
- Publication Characteristics (or Types): describe the item being indexed rather than its topic. There are 3 main categories:
    - Publication Components, e.g., English Abstract
    - Publication Formats, e.g., Lectures, Letter
    - Study Characteristics, e.g., Clinical Trial, Twin Study
- Supplementary Concept Records: allow the indexing/searching of chemicals (in the broadest sense) which are not MeSH headings

Philosophy of indexing

- The content and format of each item are fully and adequately described
- The most specific vocabulary terms are used
- The indexer's job is only to index, not to interpret, evaluate or diagnose

Finding Terms to Describe Subject Content

NLM's MEDLINE indexers use the MeSH Browser, an online vocabulary look-up aid with virtually complete MeSH records, to find the term that best describes the concept to be indexed. They view the full record: scope note, annotation, See Also terms, etc., for hints on indexing.

 Note: Searchers should use the MeSH Database to build a search strategy for PubMed or other Entrez Databases using MeSH vocabulary and to discover helpful information about a MeSH term. Use the MeSH Browser for additional information.

Indexers use the hierarchy to find the most specific MeSH heading.

Example:	"The liver disease, chronic hepatitis B" is indexed as: **Hepatitis B, Chronic** rather than Liver Diseases or Hepatitis or Hepatitis B

Indexers use more than one heading if a single heading does not cover the concept (see Coordination).

Example:	"Mucinous adenocarcinoma of the ovary" is indexed as: **Adenocarcinoma, Mucinous** **Ovarian Neoplasms**

If an exact heading does not exist, indexers use the MeSH Browser and the hierarchy to find the most specific heading available.

Example:	"Cranial radiation therapy" is indexed as **Cranial Irradiation**

which is under the Radiotherapy branch of the MeSH tree

Major Topics

- Asterisks on MeSH headings and subheadings (e.g., Wound Healing/radiation effects*) designate that they are the major topics of the article, usually obtained from the title and/or statement of purpose
- Non-major (non-asterisked) headings and subheadings are usually additional topics substantively discussed within the article, terms added to qualify a major topic, or check tags. Check tags are never major topics.
- The only indexed MEDLINE citations without an asterisked heading are some biographies in which the subject's name may be considered the only major point.
- Supplementary concept headings cannot be asterisked or carry subheadings. However, they are mapped to a MeSH heading that is automatically added to the citation. The mapped MeSH heading can carry those attributes.

Gene Links

Citations for articles in which the function of one or more genes and/or proteins is a major topic are linked to entries in the Entrez Gene database.

Coordination

Most concepts cannot be adequately described with a single MeSH term. Coordination is the use of a combination of the appropriate MeSH headings, subheadings, and check tags to index a concept as specifically as possible.

Coordination can be accomplished as follows:

Using a subheading to describe a specific aspect of a MeSH heading:

Example: "Radiographic imaging of a lung tumor" is indexed as:

Lung Neoplasms/radiography

Using two MeSH headings:

Example: "The medical staff in teaching hospitals" is indexed as:

Medical Staff, Hospital

Hospitals, Teaching

Coordinating subheadings on two or more headings:

/drug therapy on a disease term usually requires /therapeutic use on one or more drug terms.

Example: "Treatment of HIV infections with HIV protease inhibitors" is indexed as:

HIV Infections/drug therapy
HIV Protease Inhibitors/therapeutic use

/metabolism on an endogenous compound may require /metabolism on an organ, an organism, and/or a disease term.

Example: "Inability to metabolize copper" (as in Wilson's Disease) is indexed as:

Copper/metabolism
Liver/metabolism
Hepatolenticular degeneration/metabolism

/pathology on a disease term often requires /pathology on an organ term.

Example: **Myopia/pathology**
Cornea/pathology

Coordinating a major (asterisked) MeSH heading with an non-major (non-asterisked) MeSH heading that further describes the concept:

Example: "Glycosylation of ceramides" is indexed as:

Ceramides/*metabolism	(indicates that metabolism of ceramides is the main point of the article)
Glycosylation	(specifies the metabolic process)

Coordinating a heading with one or more check tags:

Example: "Preventing pre-eclampsia" in women is indexed as:

Pre-eclampsia/prevention & control
Pregnancy (check tag)
Female (check tag)
Humans (check tag)

Indexing a "pre-coordinated" MeSH heading which combines two concepts into one:

Example: "Staphylococcal pneumonia" is indexed as:

Pneumonia, Staphylococcal

Rather than:

Pneumonia (or Pneumonia, Bacterial)
Staphylococcal Infections

Depth of Indexing and the Rule of Three

- For most research articles, every concept substantively discussed should be covered by at least one MeSH heading
- Even negative findings are indexed if discussed
- For articles discussing many subjects, general headings (but as specific as possible) may be used to group related concepts rather than indexing them individually
- Some items, such as review articles, letters, and editorials, are indexed "non-depth," meaning:
    - All major topics are indexed
    - All minor concepts required for complete coordination of the major topics are indexed.

Rule of Three

- More than 3 major concepts: If more than 3 related topics are each presented as a major topic of the article, the more general MeSH heading under which they are all treed is usually designated as the major topic. The specific headings are usually indexed, but not as major topics.
- More than 3 non-major concepts: If more than 3 related concepts are discussed in an article but are not a major topic, the more general MeSH heading under which they are all treed is usually indexed. The specific headings usually are not indexed.

Note: MeSH Vocabulary present on MEDLINE citations reflect the rules in place at the time of indexing. These rules change over time. This may explain unexpected retrieval in PubMed.

Example of MeSH Indexing for a Specific Article

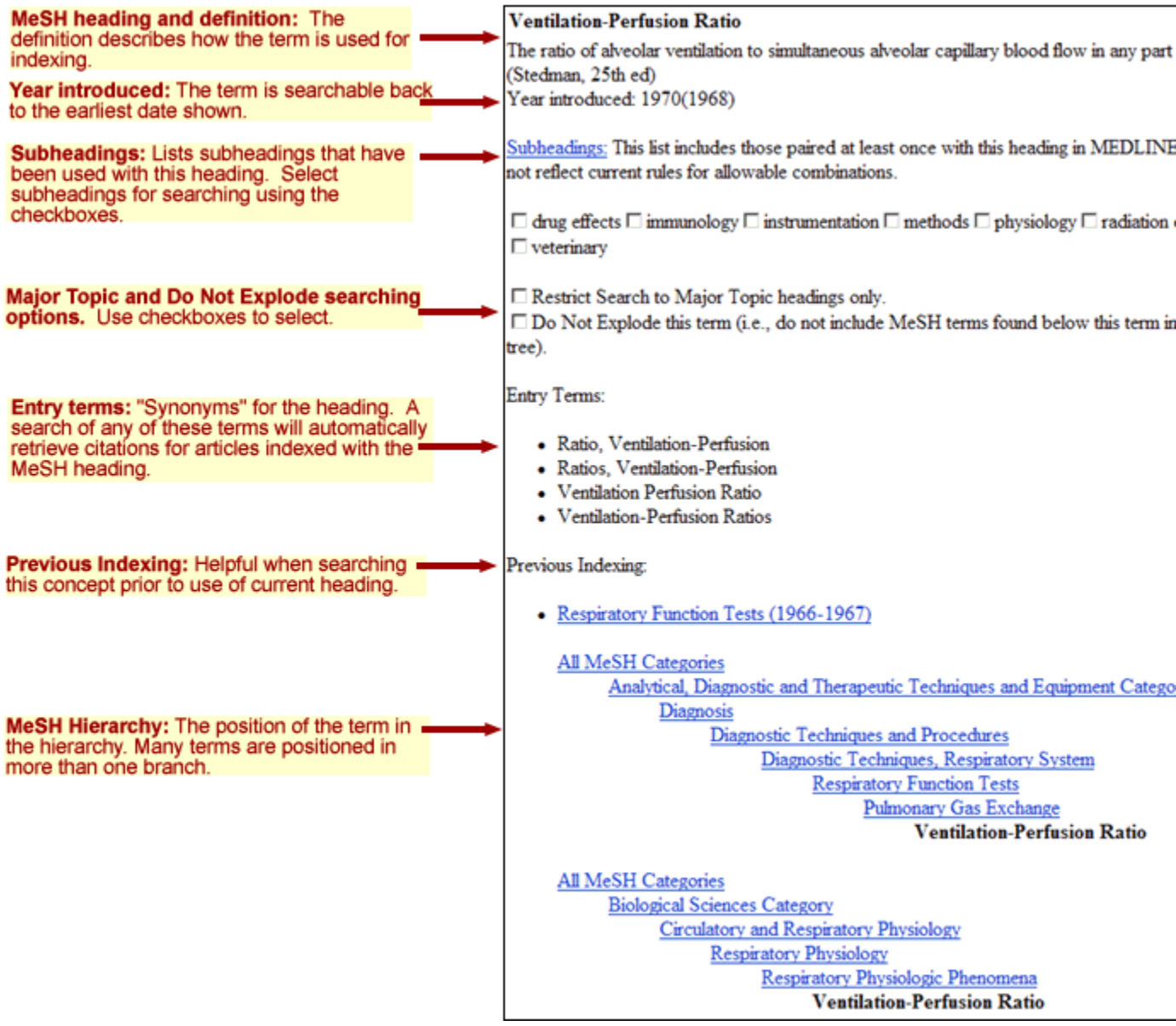
	Article Title: The role of coenzyme Q10 in heart failure. **Abstract:** OBJECTIVE: To review the clinical data demonstrating the safety and efficacy of coenzyme Q10 (CoQ10) in heart failure (HF). DATA SOURCES: Pertinent literature was identified through MEDLINE (1966-January 2005) using the search terms coenzyme Q10, heart failure, antioxidants, and oxidative stress. Only articles written in the English language and evaluating human subjects were used. DATA SYNTHESIS: HF impairs the ability of the heart to maintain its normal cardiac output. Following an initial insult, cardiac remodeling ensues, resulting in left ventricular dilation and hypertrophy. Oxidative stress is also increased, while CoQ10 levels are decreased in patients with HF. This has led to the hypothesis that CoQ10, an antioxidant, may decrease oxidative stress, impair remodeling, and improve cardiac function. CONCLUSIONS: Large, well-designed studies on this topic are lacking. The limited data from well-designed trials indicate there may be some minor benefits with CoQ10 therapy in ejection fraction and end diastolic volume. CoQ10 therapy has been shown to be relatively safe with a low incidence of adverse effects. **Publication Types:** Review **MeSH Terms:** Antioxidants/therapeutic use* Heart Failure/drug therapy* Heart Failure/pathology Heart Failure/physiopathology Humans Oxidative Stress/drug effects Ubiquinone/analogs & derivatives* Ubiquinone/therapeutic use Ventricular Remodeling/drug effects **Substances:** Antioxidants Ubiquinone coenzyme Q10
Asterisks indicate a major topic of article	
Substances include supplementary concept terms and MeSH terms, which are repeated above.	

The MeSH® Database

The MeSH Database available from PubMed's sidebar allows you to:

- locate and select MeSH terms (Headings, Subheadings, Supplementary Concept Records, and Publication Types) for use in PubMed searches
- see the definition and other helpful information for a MeSH term
- see the position of MeSH terms in the hierarchy
- select MeSH heading/subheading combinations to build a PubMed search

An example of a full MeSH Database record (see text version):



Note: Searching with MeSH subject terms excludes in process and publisher-supplied citations, as well as other PubMed citations that are not indexed for MEDLINE (e.g., citations that are out of scope for MEDLINE, such as a volcanology article in *Science*). These records do not (or do not yet) include MeSH subject terms.

MeSH Database Quick Tours

To learn more about using the MeSH database when searching PubMed, try the following Quick Tours (Adobe Flash® format):

- Searching with the MeSH Database
- Combining MeSH Terms Using the MeSH Database
- Applying Subheadings and Other Features of the MeSH Database

Searching PubMed using MeSH® Field Tags

Untagged terms that are entered in the PubMed search box are automatically mapped to the MeSH vocabulary when a match is found. However, you may choose to search the MeSH headings specifically using search field tags (also called qualifiers).

Note: Searching with MeSH subject terms excludes in process and publisher-supplied citations and others that do not (or do not yet) include MeSH subject terms.

MeSH headings [MH]

- MeSH headings can be qualified using two search field tags:

  [mh] to search a MeSH heading
  [majr] to search a MeSH heading that is a major topic of an article

  Examples: *eye [mh]*
  eye [majr]

- PubMed **automatically** searches the MeSH headings as well as the more specific terms beneath that heading in the MeSH hierarchy. This is known as the **explosion** feature.

  Example: *eye [mh]* will retrieve citations indexed to:

  Eye
  Eyebrows
  Eyelids
  Anterior Eye Segment
  Eyelids
  Lacrimal Apparatus
  Oculomotor Muscles
  Pigment Epithelium of Eye
  Retina
  Sclera
  Uvea

Vitreous Body

and all more specific terms underneath these in the hierarchy.

- To turn off automatic explosion of MeSH headings, use one of the following tags:

  [mh:noexp] or [majr:noexp]

  Example: *eye [mh:noexp]*

This retrieves citations indexed with the term, Eye – but not necessarily indexed with the terms beneath that heading in the MeSH hierarchy. Remember that indexers use the most specific MeSH heading to describe the subject of a journal article.

Subheadings [SH]

- A list of subheadings can be found in PubMed's Help.
- MeSH specialists and indexers determine which subheadings may be used in combination with each MeSH heading. These combinations may change over time.
- For searching, attach subheadings to MeSH headings using the format: MeSH heading/subheading.
- Two-letter abbreviations for subheadings or the full subheading name may be used.

Examples: *thromboembolism/pc [mh]*
thromboembolism/prevention and control [mh]
toes/in [majr]
toes/injuries [majr]

- Only one subheading may be attached to a MeSH heading at a time. To attach multiple subheadings, combine each MeSH/subheading combination with the OR Boolean operator or use the MeSH Database, which allows multiple subheadings to be selected.

Example: To search for citations where the main topics are about the prevention and control or diagnosis of thromboembolism, enter:

thromboembolism/pc [majr] OR thromboembolism/di [majr]

- For a MeSH/subheading combination, PubMed always explodes the MeSH term ***and*** also searches the subheading and its grouping if there is one. You cannot explode a MeSH heading and not also explode the attached subheading.

Example: *polysaccharides/ae*

Retrieves citations indexed with any of the polysaccharides with any of these subheadings:

ae (adverse effects)
po (poisoning)
to (toxicity)

Note: *polysaccharides/ae [mh:noexp]* turns off both the MeSH heading explosion and the subheading explosion.

Unattached subheadings

- You may also choose to search a MeSH heading and a subheading combination using the AND Boolean operator and the subheading field tag of [sh]. This may be done when you want to search for a subheading concept that cannot "legally" be attached to the MeSH heading you are also searching.

Example: *hypertension [mh] AND toxicity [sh]*

- To **turn off the subheading grouping**, use the tag [sh:noexp]. Do this only when using an unattached subheading.

Example: *finger injuries [mh] AND surgery [sh:noexp]*

PubMed's Help includes:

Subheading list

See also:

Subheading (Qualifier) Hierarchy

Indexing for Chemicals and Drugs

- The D branch of the MeSH vocabulary tree is for Chemicals and Drugs.
- Frequently, chemicals that are not currently MeSH headings need to be used for indexing. To enhance coverage of chemicals, Supplementary Concept Records (SCR) are available.
- Every SCR is mapped to a MeSH heading which is automatically added to the citation when an indexer chooses an SCR to describe an article. These mapped terms can be found in the MeSH Database:

Example: *coenzyme Q10 [Substance Name]*

Heading Mapped to:
Ubiquinone/analogs and derivatives

- Drugs and chemicals may also be searched in the MeSH database using their Registry Numbers or Enzyme Commission numbers.

Example: *86386-73-4 [rn]*

Note: Not all records for chemicals in the MeSH database have registry number data, so retrieval may be incomplete. If a search using a registry number is unsuccessful, search on the substance name instead. Classes or groups of drugs or chemicals, e.g., benzenes, are not assigned specific Registry Numbers.

Pharmacological Action Terms

Most chemical and drug MeSH headings and Supplementary Concept Records (SCRs) are assigned one or more headings that describe the pharmacological action(s) (PA) of the substance.

- Since 1996, NLM indexers add the appropriate pharmacological action MeSH heading as well as the specific chemical MeSH heading to a citation when the action of the chemical is being discussed in the article.

Example: Here are the pharmacological actions established for the MeSH Heading, Aspirin:

MeSH Heading	**Aspirin**
Pharmacological Action	Anti-Inflammatory Agents, Non-Steroidal
Pharmacological Action	Cyclooxygenase Inhibitors
Pharmacological Action	Fibrinolytic Agents
Pharmacological Action	Platelet Aggregation Inhibitors

- An article indexed since 1996 that discusses **aspirin used as an anti-inflammatory agent** is indexed with:

  Aspirin
  Anti-Inflammatory Agents, Non-Steroidal

- An article indexed since 1996 that discusses **aspirin used to inhibit blood clotting** is indexed with:

Aspirin
Platelet Aggregation Inhibitors

- To search for articles indexed since 1996 about a particular action of a substance, search the substance term AND the Pharmacological Action *as a MeSH heading*.

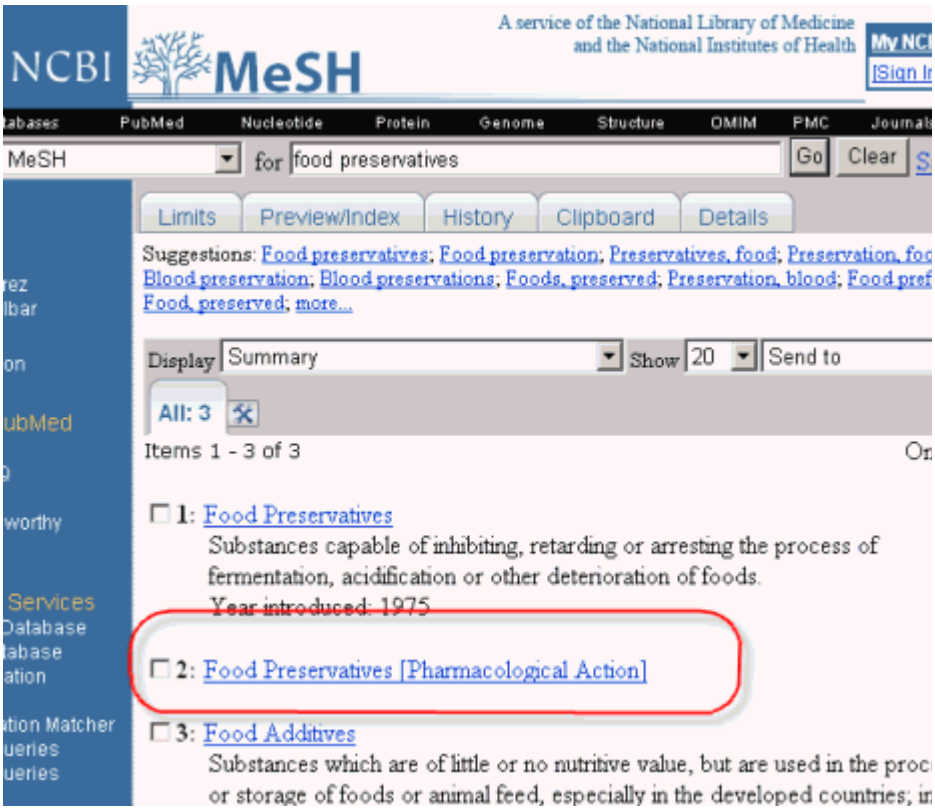
Example: *aspirin [mh] AND fibrinolytic agents [mh]*

Pharmacological Action Lists

In addition to their use as MeSH headings for indexing, pharmacological action terms have separate records in the MeSH database that list the substances that are assigned to that action:

Type the Pharmacological Action term in MeSH Database search box.

Click on term labeled [Pharmacological Action].



The Full format displays the substance terms that have been assigned the Pharmacological Action term, **Food Preservatives:**

Food Preservatives [Pharmacological Action]

- allyl isothiocyanate *(Substance Name)*
- Benzoic Acid *(MeSH Term)*
- juglone *(Substance Name)*
- Nisin *(MeSH Term)*
- octyl gallate *(Substance Name)*
- Parabens *(MeSH Term)*

- purpurogallin *(Substance Name)*
- Sodium Benzoate *(MeSH Term)*
- Sodium Nitrite *(MeSH Term)*
- Sorbic Acid *(MeSH Term)*


Note: In the MeSH Database, items marked *Substance Name* are Supplementary Concept Records. When searching PubMed, a search of substance name [nm] will retrieve ***both*** Supplementary Concepts ***and*** MeSH terms.

- Use of a term with the Pharmacologic Action search tag [pa] instructs PubMed to OR together terms from a list made up of a PA term and the drug/substance terms known to have that action.
- MeSH terms on the list are searched with the no explode specification, [mh:noexp], so as not to include narrower terms that might not share the pharmacological action.
- Use this search method when you want to include retrieval for all substance terms from the pharmacological action list.

Example: *beverages [mh] AND food preservatives [pa]*

- Remember, to search for articles about one particular action of a substance, search the Pharmacological Action *as a MeSH heading*, so as not to include all of the other actions of that substance. Retrieval of citations indexed before 1996 is likely to be incomplete.

Note: If you enter a MeSH term that happens to be a PA term without using a search tag, PubMed will search the term as [mh] OR [pa] OR [tw].

Searching for Chemicals and Drugs

Use the Substance Name [nm] in PubMed for comprehensive chemical searching, whether the substance you are looking for is a MeSH heading or a Supplementary Concept Record (SCR).

Example: *ketorolac [nm]*

Note: Use text word [tw] searching or use Previous Indexing headings found on MeSH heading records in the MeSH Database for the years prior to the introduction of the chemical into the controlled vocabulary.

Note:

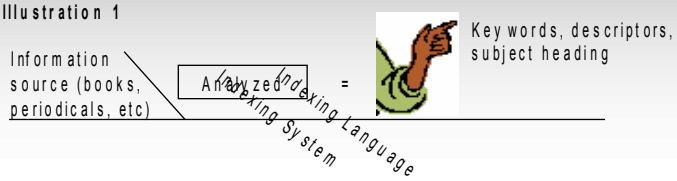
In the MeSH Database:	In a MEDLINE citation:
Chemicals, Drugs and Compounds listed as **[Substance Name]** are Supplementary Concepts (SCRs) only	Chemicals, Drugs and Compounds listed in a citation under **Substances** can be either MeSH terms (in which case, they are duplicated in the MeSH Term list) OR Supplementary Concepts (SCRs) Search ***all*** substances (both SCRs and MHs) in PubMed using the search tag **[nm]**

INDEXING

Indexing Defined –

… is the process of assigning a subject heading to an information source of whatever form (books, periodical, multi-media, etc.) to serve as retrieval tool for that information source being analyze.

Illustration 1



Six Different Methodologies for Indexing

- Stand-alone or Dedicated
- Embedded Indexing
- Tagging
- Keywording
- Automated
- Free-text and Weighted-text
- Controlled-Vocabulary Indexing

Standalone or Dedicated Tools

Usually used for back-of-the-book indexes, allow indexers to work from page-numbered galleys. The indexing is completely separate from the published material.

Embedded Indexing

Is the process of creating index entries electronically in a document's files. The indexer inserts the index entries as invisible text in the electronic files. Although desktop publishing packages are not the best tools for indexing, they can be used to create effective embedded indexes. For technical documents that will be updated frequently or will go on-line, indexers can create embedded indexes that will help the audience find information quickly and efficiently. Some tools that allow for embedded indexing include FrameMaker, Microsoft Word, Adobe PageMaker, and Quark Xpress. Markup languages that allow for embedding indexing include SGML, HTML, XML, and TeX/LaTeX.

Tagging

Allow indexing codes to be embedded in the electronic text after the indexing is complete. The indexer inserts numbered dummy tags in the files, and then builds the index separately. The final step uses macros to insert the indexing at each tag in the files. Many of these tools are developed-in-house to fit the publishing group's needs..

Keywording

Is used primarily in online help materials. It can be inserted as embedded coding and built into a list by the software.

Automated Indexing

A tool that now accompanies most processing software, build a concordance or a word list, from processed files. Although the manufacturers often claim these packages build indexes, the beginning stages of building an index. Usability tests of these packages have shown that the word lists omit many key ideas and phrases, and cannot fine-tune terminology for easy retrieval, or build the needed hierarchies of ideas that professional indexing can.

Free-text and weighted-text

Searching tools are not discussed in these pages, but are aspects of information retrieval that indexers are very interested in.

Controlled-Vocabulary Indexing

Involves the used of Sear's list, classification Table to accurately assign subject heading to a document. It also enables the professional indexer to fine-tune terminologies, build needed hierarchies of ideas and to go to the most specific term or phrase to describe a document.

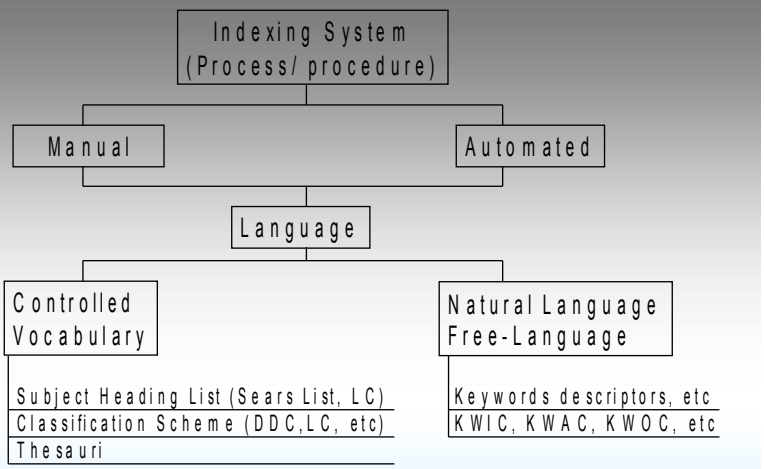


Fig. 1.0 Indexing Methods

Indexing language

The language used to described the subject or other aspect of information or document in an index. It is a system for identifying or naming subjects contained in a document.

Elements:

- **Vocabulary** - the terms used in the indexing of concepts.
- **Syntax** - the way the words are put together to form index terms. It involves coordination to identify a particular concepts.
- **Semantics** – meaning of words this indicates the class relations among index terms. Semantics relationship is categorized into:

  Equivalence Relationship – It is possible to have more than on term for the same concept such as synonyms, preferred spelling, acronyms, etc.

  Hierarchical relationship – a genus-species relationship which indicated class inclusion such as: Horticulture-Ornamental-Orchids, etc.

  Affinitive/ Associative relationship –indicates through related terms such as: Land and Sea, Birds and Sky, etc.

Types of Indexing Languages

1. **Controlled Vocabulary:**

   An authority list which enable an indexer to establish a standard description for each concept and used that description each time it is appropriate, in order to:

   - Control the synonym by choosing one term as the standard form.
   - Make distinctions between homographs.

     ex. Ball (Basketball/ Volleyball) Ball (J-S Prom)
   - Bring or link together those terms whose meanings are closely related

     ex. Beverages and Drinks

Examples of controlled vocabulary

- Subject Heading List – Sears List of Subject Heading, LC, DDC
- Classification schemes – DDC, LCC, UDC, Colon Classification
- Thesauri

Types of Indexing Languages

2. Natural/ Free Language

Uses index terms derived from the documents. There are no limits to terms that will be used as index entries. It allows opportunities for recall since it provides more access points; results to greater redundancy; and reduces precision.

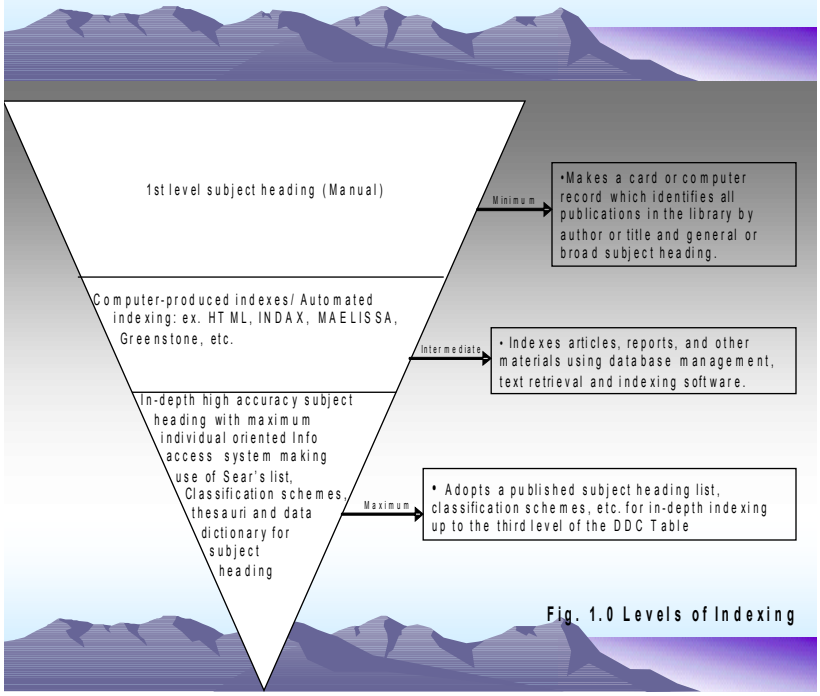


Fig. 1.0 Levels of Indexing

Process/ Procedure in Indexing:

1. **Bibliographic description** – follow given template/ worksheet of software used
2. **Analysis of contents** – aim at in-depth indexing
   - Examine title, abstract, table of contents, text and reference section of the item.
   - Relate the content to the user
   - Subject determination. Subjects represented in the work should be identified in the work, making up a list of possible descriptors.
   - Conversion into the indexing language. Translate indexable concepts into the standard index terminology.
   - Prepaid index entries.
3. **Production of index entries** – hard copy/ soft copy

Some Principles of Indexing:

- Correctness (Accuracy) – correct spelling and correct presentation of information in the index
- Convention (Uniformity/ consistency) – because indexing systems vary, indexers must adopt uniformity in indexing
- Completeness (Complete bibliographic data) -
- Clarity (Specific and concise subject heading) –
- Consideration (In the language of the user) –

Some Requisites of a good indexer

- Subject knowledge
- Knowledge of user needs
- Experience
- Concentration
- Reading ability
- Comprehension

- **Width / Extensive Exhaustiveness (*high recall, high precision*)** – can be measured by the number of index terms and number of items indexed in a document
  - Refers to the number of topics covered or degree of details covered for a particular topic in indexing a document
  - The degree to which all the separate subjects in a given document are noted in the indexing activity and converted into the language of the indexing system
- **Specificity (*low recall, high precision*)** – the degree to which the descriptor matches the subject concept
  - Also called "depth indexing"
  - means the power of the index to precisely define the topic
  - Deep specificity gives low recall and high precision. Scant specificity results in high recall and low precision
- **Consistency** – extent to which agreement exist on the terms to be used to index the documents.

TWENTY RULES IN INDEXING

1. Index everything useful in the source materials
2. Include all index entries in one alphabetical sequence
3. Choose popular headings, with references from their specific equivalents except where specialist audience is addressed.
4. Be consistent in choosing one form of spelling. Use a standard dictionary.
5. Choose the most specific headings which describe the items indexed.
6. Be consistent in the use of singular form or plural terms
7. Combine the word and the action which describes it where it is useful and possible
8. Invert headings, when necessary to bring significant word to the fore.
9. Check for synonyms and make suitable references.
10. Check for antonyms and combine where suitable.

TWENTY RULES IN INDEXING

11. Where words of the same spelling represent different meanings, include identifying phrase in brackets.
12. Where possible, give full name of persons quoted.
13. Omit the name of the country of a government or department
14. Use capitals for all proper names.
15. Make references from main subdivisions of these subjects.
16. Subdivide alphabetically by aspects wherever possible.
17. In the case of historical or biographical works, substitute chronological for alphabetical subdivisions.
18. Spell out symbols and abbreviations
19. Avoid the use of bold type wherever possible.
20. If references are made to paragraph numbers and to page numbers, include a note to this effect at the foot of every page of the index.