

machine, such as the entrance requirements for a specific graduate school, or the location of articles published in the *journal of distance Education*, then you would begin your search at a multipurpose *search engine* (such as Google or Alta Vista). If, however, you were interested in more general information, such as what schools offer graduate education programs or a broad source of information about distance education, then your most effective search tool would be a hierarchical *subject guide* (such as Yahoo!).

In hierarchical directories, or subject guides (such as Yahoo!), you would scroll down through hierarchical menus until you arrive at the site you are looking for. In simple operational terms, subject guides are organized around directories, which are created and populated by humans who have reviewed the listed sites and determined the particular subject headings under which they are most usefully classified. The benefit of subject guides over search engines is that the searcher does not have to wade through hundreds of documents to find relevant information. Directories provide the searcher with a straightforward, hierarchical means of information retrieval on the Web. However, this is precisely the reason that subject guides are not used with the frequency that search engines are. While subject guides can save the searcher much time by selecting only certain Web sites for each category, this also means that many sites are excluded. It is not possible for any group of people to accurately and continuously categorize and re-categorize all information available on the Net and accurately assign it to all possible relevant categories. As such, subject guides are not bias-free, because someone has made a decision about which resources will be included in the guide and under what categories.

On the other hand, search engines (such as Google, Alta Vista, Infoseek, HotBot, Excite, Lycos) use each word in the document and all meta tags contained in the page's introduction to create very large databases. These databases are developed and maintained automatically by software robots commonly known as *search bots*—or just bots. The search algorithm links your request with the titles and phrases found on the Web sites and stored in the databases. In simple operational terms, search engines work with robots that index the contents of nearly every document on the publicly available Web. Then the contents are entered into databases on very large and fast search engines. When you enter a query at a search engine Web site, your input is checked against the search engine's keyword indices. The best matches are then returned to you as hits. Most search engines use search-term frequency as a primary way of determining the order in which the hits are organized, by listing first the document with the most uses of the keyword. If, for example, the word or one of the words in your query appears numerous times in a Web document, it is reasonable to assume that the document will likely turn up near the beginning of the search engine's list. Some search engines are designed to search for both the frequency and the positioning of keywords to determine relevancy, reasoning that if the keywords appear early in the document or in the headers, the document is more likely to be on target.

Finally, some of the better search engines use a technique (such as that developed by Google.com) that tracks which sites are actually visited from the hundreds that are returned as hits and places these "consumer choice" sites higher on the list with each subsequent search. Thus, these systems become more accurate and useful with use and are especially useful for finding often-searched sites and terms. The search bots constantly update the indexes, usually visiting popular sites every few weeks. It should also