INFORMATION SEARCH AND RETRIEVAL

INTRODUCTION

Since humans first began to use tools, there has been a need to convey technological information. The earliest transmission of such information was by word of mouth. For thousands of years instructions in the making and use of tools were handed down from generation to generation within families or small communal groups. The invention of writing, in about 3400 BC, expanded the audience for technological information to those who could understand written symbols (1). Indeed, much writing in the ancient world concerned the transmission of technical information. The invention of paper by the Chinese further expedited communication. In the Middle Ages, scholars dispatched personal letters to announce new discoveries. Echoes of this practice can still be found in publications such as *Electronics Letters*, which is dedicated to the rapid communication of research results. The invention of the printing press in 1453 created the potential for distribution of information to a mass audience. The partial realization of this potential was a factor in the scientific revolution two centuries later. Hooke, Newton, Leibnitz, and their colleagues published books and broadsheets describing their theories and experimental results. These publications stimulated thought among the small but ever-growing body of the scientific and technologically curious. This outpouring of scientific activity in the late seventeenth century led to the establishment of the Royal Society, the first institution dedicated to the encouragement of scientific research. At first, the Society was a discussion group, a sounding board for new ideas. Gradually, it developed more formal means for the communication of new scientific knowledge. By the early nineteenth century, when Michael Faraday began the serious study of electricity and its properties, the Royal Society was publishing the first modern scientific journals. The defining features of these publications were both a dedication to communication of new scientific knowledge and a rigorous critical review of such communication. Within 50 years, early professional societies in engineering had adopted the model of the scientific, or scholarly, journal. Other media have also had prominence in the communication of knowledge within electrical engineering. Patents were probably the primary source of communication among electrical engineers and inventors in the nineteenth and early twentieth centuries. Edison, Tesla, Westinghouse, and many of their engineering contemporaries announced the results of their work almost exclusively via the means of patents. With the advent of computer technology, the dissemination of information has undergone another revolution. The defining characteristics of this revolution have been increases in storage capacity, processing power, and speed. Engineers can now transmit masses of data with the touch of a few keystrokes. These developments promise to produce new, and heretofore unknown, forms of scientific and technological communication.

FORMS OF COMMUNICATION IN ELECTRICAL ENGINEERING

Scholarly Journals

The medium of the scholarly journal is still paramount as a means of communication among scientists and engineers. The discipline of electrical and electronics engineering is no exception. A common feature of most scholarly journals is that they are peer-reviewed. Peer reviewing is a process that has evolved to lend assurance of scientific accuracy. Articles submitted to these journals are sent to reputable, impartial practitioners in the discipline. These reviewers then check the articles submitted to them for accuracy of research methods, relevance to the discipline, and plausibility of results. Normally, this process is one of blind review, meaning that the identities of authors and reviewers remain unknown. Scholarly journals are often published by established professional societies in each discipline. In electrical engineering, two such societies, the Institute of Electrical and Electronics Engineers (IEEE) and its British counterpart, the Institution of Electrical Engineers (IEE), publish the most prominent English language journals. The IEEE is particularly well known for its *Transactions* series of publications, which are peer-reviewed journals devoted to specific aspects of electrical engineering. The IEE publishes a number of scholarly journals as well. Chief among these journals are: *IEE Proceedings*, *IEE Review*, and *Electronics Letters*. The American Institute of Physics also issues journals, which are of interest to electrical engineers, including: *Journal of Applied Physics, Physics of the Solid State, Semiconductors,* and *Technical Physics*. In addition, various commercial publishers, such as John Wiley & Sons, Springer Verlag, Elsevier Scientific, and McGraw-Hill, publish scholarly journals in electrical engineering.

Some scholarly journals are now available through a process known as open access. In this model costs of distribution are kept to a minimum by mounting articles directly on the Internet. Costs of production are typically handled through page charges to authors rather than subscription fees. For more information about open access see the Scholarly Publishing & Academic Resources Coalition (SPARC) Web site at: http://www.arl.org/sparc/.

Other publications concentrate on the news or current events in electrical and electronics engineering. A good example of this type of publication is the *IEEE Spectrum*. Although these titles normally do not publish research results, they would include reports on trends in research.

Conference Proceedings

Research results and other scholarly information are also regularly disseminated through professional and scholarly conferences. This process normally works in two ways. First, papers are presented orally to the conference attendees. Then, these papers are published in the proceedings of the conference. Many conferences are sponsored and organized by professional societies and are often held annually or at other regular intervals. Conferences are also organized by universities, institutes, government agencies, and international organizations. Conferences provide opportunities for electrical and electronics engineers to inter-

act with colleagues from other parts of the world. Indeed, they sometimes provide unique occasions for communicating the ideas and views of engineers who, because of language, lack of communication infrastructure, political climate, or other reasons, may be isolated from professionals and scholars with similar interests. Copies of the written record, or proceedings, of conferences are usually provided to attendees as part of the registration package. In addition, proceedings are normally published by the sponsor of the conference or by a commercial publisher. Published proceedings are offered in limited runs and are usually purchased exclusively by libraries. Research findings and other information presented at conferences can be subject to peer review. Often, however, more preliminary findings are presented at conferences, preparatory to later publication in scholarly journals.

Standards

Standards are a particularly important form of communication in electrical engineering. Standardization in engineering can be described as: "The process of establishing by common agreement engineering criteria, terms, principles, practices, materials, items, processes, equipment parts, and components (2)." In the United States, the most prominent standardization agency for electrical engineering is the IEEE. Committees composed of members develop standards within the IEEE. These groups normally compose a draft standard, which is then submitted to the IEEE Standards Board for final approval. IEEE standards are now also available online. The Underwriters Laboratories (UL) issues standards for electrical products and appliances. Other societies issuing important standards in electrical engineering include: Comite Consultatif International Telegraphique et Telephonique (CCITT) and the International Electrotechnical Commission (IEC). General engineering standards, which often affect electrical and electronics engineering, are issued in the United States by the American National Standards Institute (ANSI) and internationally by the International Organization for Standardization (ISO).

Patents

Patents are an essential part of the information infrastructure in electrical engineering. The primary purpose of patents is to protect the rights of inventors by granting them temporary monopolies over their inventions. Patents are issued by national governments, which means that differences can exist from country to country in terms of format, procedure, and scope of protection. Patents also serve as a rich source of information in engineering. In applying for a patent, inventors are required to state all they know of the invention. In most cases, this information is not available elsewhere. In fact, it has been shown that up to 90% of patent documents are never reported in the journal literature or in other media (3). Patents are usually available to interested parties through government patent offices and, in some cases, in selected libraries. The United States, along with most other patent-granting countries, issues a patent gazette that provides an extensive abstract of each patent and, often, drawings. A drawback is that these publications are usually arranged by an arbitrary numbering system. To complete a proper search of this literature requires a patent index, most of which are available in electronic form. The United States Patent Office offers online searching via the World Wide Web at: http://www.uspto.gov. The Patent Office also provides full-text of patents. To view the images in patents, however, a Tagged Image File Format (TIFF) plug-in must be installed, which can be done by following the links on the Patent Office search page. Patent databases are available commercially as well. Two of the best known are the *Derwent Patent Citation Index* and the *Derwent World Patents Index*.

Government Information

Government reports and regulations are a valuable and necessary source of information in electrical engineering. Various government entities carry out basic and applied research. Often, this research is only reported in government publications. Laws and regulations govern all types of electrical engineering practice. Good examples are those dealing with the disposal of hazardous waste. Government agencies and commissions are often instrumental, through their reports and other publications, in setting national policies, which affect the practice of engineering. In the United States, this data, and other information from the national government, is provided to the public through a system of depository libraries. United States government documents are increasingly available on the Internet as well. These documents can often be found by entering the website of the agency that issued them (i.e., Department of Energy, Environmental Protection Agency). Another option is to access the Government Printing Office website at: http://www.gpo.gov. Other governments, particularly in developed countries, have somewhat similar programs, usually coordinated through national libraries. Governments also provide funding for engineering research carried out in universities, institutes, and private engineering firms. Reports based on this research are, in many cases, subsequently published by government agencies or by private firms with exclusive rights to its publication. In the United States, the National Technical Information Service (NTIS) is a private firm that acts as a clearinghouse for this type of information. The NTIS maintains a database to provide access to these reports and sells them to interested parties. Not all government reports are available to engineers and other members of the public. Most governments have established classification schemes for restricting access to sensitive documents. Often, but not always, these documents relate to national defense.

Gray Literature

The report literature, or gray literature, is a valuable, yet elusive, source of information for electrical engineers. The gray literature refers to internal reports, memoranda, unpublished manuscripts, notes, and other forms of information that are rarely indexed or rarely find their way into library collections. Clues to the existence of such information can sometimes be found in the citations of journal articles or in specialized bibliographies. Obtaining these reports, however, can be difficult. Requests to one or more of the

authors can be the most efficient way to obtain this material. Requests to corporate archives and documentation centers can also be successful. In some cases, this kind of information can be located on the Internet. Complicating the situation is the fact that these reports are often proprietary and, thus, not available to the public.

BIBLIOGRAPHIC TOOLS

Large amounts of information, however valuable, are not much use without the tools designed to locate and use it efficiently. In electrical engineering and electronics, a number of such tools exist. These tools usually take the form of databases and indexes that provide access through a number of different search parameters, such as subject, author, and title, to the literature of disciplines, subdisciplines, and topic areas. Indexing sources can assume a variety of different formats and can appear in various media. Paramount among these media are print, CD-ROM, and online. Common to all indexing sources is the concept of the record. A record, in indexing terms, is a unit of information that describes a specific published entity (although some indexing records describe unpublished manuscripts as well). Such entities can assume a number of different forms: articles, conference proceeding papers, books, government reports, Internet publications, and others. All records provide certain basic pieces of information about published entities, such as author(s), title, subject or keyword, and, if describing an article, the source periodical where the article can be found. Indexing sources that provide access to the literature in scholarly journals and conference proceedings often provide abstracts as well. Abstracts provide a description of a scholarly paper. Normally, this description is a paragraph in length, but it can be more extensive.

Electronic Indexes

Increasingly, bibliographic databases in electronic media have become the most heavily used tools for accessing the literature of electrical engineering. The reasons for this development are many; databases provide almost instantaneous access to search results across a number of years, whereas print indexes rely on laborious search processes, which cover one year or, at best, a short span of years. A further advantage of databases is that, with the advent of online access through the Internet, searches can be performed from offices and homes, maximizing efficient use of libraries. To provide access to their records, most databases use a form of boolean logic in which sets of record addresses are established that conform to search parameters provided by the user. The usual pattern is that a set, containing "x" number of records, is established, which is then modified by further sets. For example, a search of the term "semiconductors" might be modified by "gallium arsenide." The three main boolean connectors used in bibliographic searching are: "and," "or," and "not." "And" is used to form sets in which each record contains all subject terms entered via a specific search statement. "Or" forms sets in which alternate terms are included, such as "Very Large Scale Integration" or "VLSI." The purpose of the "not" connector is to exclude records that mention the indicated subject terms. Some databases also use proximity connectors. These produce records in which subject terms must occur within a defined proximity to each other, such as a sentence, a paragraph, or within a specific number of words. Another standard tool for database searching is truncation. This tool is a variation on the "or" connector in which the root of a subject term is defined and records containing all variations of that root are retrieved. For example, a truncation search on "circuit*" would retrieve records that contain "circuit," "circuits," or "circuitry."

A more refined method of searching is to define the search to a specific field, such as subject, author, or title. The advantage of this method is that an undefined search might produce hundreds or perhaps thousands of records, whereas a search limited to specific fields will tend to produce not only a smaller number of records, but records that are more focused to the result the searcher wishes to obtain. For even greater precision in subject field searches, many databases also employ controlled vocabularies of subject terms, which lends precision by establishing authority control for subject terms. Authority control means that one master subject term is assigned to records on appropriate topics, in place of several variants. This practice tends both to increase the number of records retrieved by a search and to focus the scope of the records retrieved more efficiently.

INSPEC is widely regarded as the premier bibliographic database in electrical and electronics engineering. The *INSPEC* database was developed by the Institution of Electrical Engineers and is still managed by the IEE. The scope of this database is: "... the worldwide literature of physics, electronics and electrical engineering, computers and control, and information technology (4)." It covers a wide range of publications in these areas; however, the primary focus is on papers published in professional journals, as well as papers presented at scholarly conferences. In these areas, *INSPEC* is particularly comprehensive, indexing a number of titles that are ignored by other abstracting services. Like other high level indexing services, such as *MEDLINE* and*PsychInfo*, *INSPEC* has its own classification scheme. Although use of this scheme is not necessary to successfully retrieve records from the database, it can lead to more precision in searching. The scheme is hierarchical. The database is divided into four sections, Section B covers Electrical Engineering and Electronics. Four levels exist within each section, each of which denotes a progressively more specific subject area (5). Records in *INSPEC* can also be retrieved using keyword, title word, author name, and other search parameters. Although *INSPEC* has a long history as a print index, it is now typically used on the Internet, usually through site licenses purchased by universities or businesses.

The IEEE offers its own bibliographic searching capability, along with electronic access to the full text of its published papers, through IEEE Xplore. This search capability is limited to journals, conference proceedings, and standards published by the IEEE. Within that scope, comprehensive searching is available. It is also possible to searching using *INSPEC* subject terms. If one has access to IEEE publications online, the search results on IEEE Xplore will link to the full text of the retrieved publications.

Another database of importance to electrical engineers is *Compendex*. This file is based on the *Engineering Index*, which has been the standard bibliographic index to the literature in engineering since the 1880s. Its scope is all areas of engineering, along with allied areas, such as construction management. Like *INSPEC*, the *Compendex* database concentrates on indexing professional journals and the papers presented at scholarly conferences. Some technical reports, books, and increasingly, Internet sites are included as well. The scope of the database is worldwide. *Compendex* is, in fact, one of the chief vehicles by which engineers in English-speaking countries may learn of research reported in other languages. Subject access to the database is accomplished through a controlled thesaurus of subject terms called the *Subject Headings for Engineering*. Keyword searching is also available, as is limiting of searches by discipline, year, and type of publication.

The *Web of Science*, a product of the Institute for Scientific Information that incorporates its *Science Citation Index*, is also of interest to electrical engineers. The unique feature of this database is that it indexes papers that have been cited in the scientific and engineering literature. This indexing tends to produce a file that focuses on more influential articles. Also, because of its orientation, *Web of Science* offers the capability of following the citation trail of particular articles or authors. This type of searching can uncover clues to the influence of particular ideas and discoveries as they reverberate through the literature. To make such searching more efficient, direct links are available both to cited papers and to papers that have cited them. Increasingly, links are also available through *Web of Science* to electronic full text of scientific papers.

For engineers involved in the production of electric power, the *DOE Information Bridge* website (http://www.osti.gov/bridge) is valuable as well. Produced by the United States Department of Energy, it includes full text of: "... unclassified scientific and technical information processed or received by the DOE Office of Scientific and Technical Information (6)." References to the published literature in all areas of energy, including alternative energy sources, are included. It includes online full text of reports published from 1994 to the present.

Another form of bibliographic database is the table of contents or contents alert service. The most prominent files in this area belong to *Current Contents*. *Current Contents* was first published by the Institute for Scientific Information in 1958 and, in the intervening years, it has become a staple in the research efforts of scientists and engineers. In its print format, *Current Contents* reproduces the title pages of prominent journals in scientific subject areas, one of which is engineering. This tool has proven to be a convenient method for scientists and engineers to learn of topics in the current literature. In recent years, *Current Contents* has become primarily available as an online electronic file. Electronic databases such as *Compendex* and *IEEE Xplore* also offer content alert services.

Subject Bibliographies

Subject bibliographies constitute an additional type of research tool. Often, these publications give a comprehensive view of the literature in specific subject areas (for example, microprocessors). Subject bibliographies may include annotations, or short descriptions, much like abstracts, of each referenced item. A variant of the subject bibliography is the guidebook, also known as a guide to the literature. These sources provide a useful way of identifying the major information sources in a given discipline.

DOCUMENT RETRIEVAL

Libraries

Once articles, books, and other documents have been identified, it is still necessary to retrieve them. Traditionally, retrieval has been best accomplished through the library at the university, institute, laboratory, or company with which an engineer is affiliated. This retrieval method remains an efficient option. The material is on site and readily accessible through subject arrangements, such as Library of Congress classification. Modern libraries have well-developed capabilities for accessing electronic databases, electronic journals, information on the Internet, as well as print resources. Most libraries also offer services for obtaining documents from other institutions. Libraries also offer comprehensive reference services, designed to address research strategies and problems. Some libraries, notably the University of Michigan and others in cooperation with Google, are creating "virtual libraries," which exist on the Internet and are not confined to a physical space (7).

Document Delivery

In recent years, additional document delivery options have become available. ProQuest and other companies act as clearinghouses for theses, dissertations, and published papers, which they then provide for a fee. Both professional societies and vendors of bibliographic databases also offer services for retrieving full-text copies of published articles. Increasingly, these services are available online through a subscription fee.

Copyright

Copyright considerations are a major concern in establishing these services. Most vendors include the copyright fee as part of the overall price for each document. In the United States, payment of copyright fees is then handled through the Copyright Clearance Center in Danvers, Massachusetts. The center acts as a clearinghouse for copyright compliance and makes payments to authors. It also maintains close ties with the U.S. Copyright Office (8). Internationally, the World Intellectual Property Organization (WIPO) works to provide copyright protection in the countries that are parties to the various treaties and conventions that govern intellectual property.

Electronic Journals

Beginning in the mid-1990s, published information began, increasingly, to appear online, particularly through the medium of the Internet. One manifestation of this phenomenon has been the electronic journal or e-journal. Al-

though e- journals can take many forms, it is most common for those in science and engineering to replicate what is available in their print versions. When articles are accessed in electronic journals, they usually appear in one of two formats, HTML (Hypertext Markup Language) or PDF. HTML is the standard programming code for the World Wide Web. Articles coded in HTML have the advantage of ready accessibility, normally within one level of the table of contents of an e-journal. The drawback of this format is that it does not permit the representation of characters, such as mathematical notation, which are outside the bounds of plain text. PDF, on the other hand, represents an exact digital copy of the printed page. The disadvantage here is that to access the PDF representation of an article, one must enter a program that allows viewing of text in that format, a process that can involve several steps. The most popular of these programs is Adobe Acrobat. This situation is mitigated by the fact that Adobe Acrobat Reader is available, free of charge, on the World Wide Web. As a result, and because of the obvious advantages of representing exact images of journal articles online, PDF has become the format of choice for electronic journals in science and engineering.

Google Scholar

In recent years, Google has emerged as the most popular Internet Search Engine. In 2004, Google unveiled a new product, Google Scholar. The aim of Google Scholar is to provide access to the academic and professional literature (9). Unlike traditional indexing sources, Google Scholar does not use a controlled vocabulary but uses Google's spider software technology to discover and make use of freely available metadata, or keywords placed in certain HTML codes. As the metadata and the records come from many different sources, it tends to produce an unstructured set of results. However, because Google Scholar is available at no cost and other indexes of the academic literature usually require purchase of a subscription license that can cost thousands of dollars, Google Scholar has become popular among professionals and students, particularly those who do not have access to the resources of large research universities.

Internet Sites

The Internet opens electronic possibilities that are unknown in other media. E-mail has transformed communication, first among scholars and now throughout the general population (10). On the World Wide Web, the changes have been even more striking. Data sets are now routinely linked to research articles. Links are provided directly to cited works. Photographs and other illustrations can be included with articles or drafts and transmitted with ease. Video can be included to more fully illustrate important concepts. Blogs (Weblogs) and RSS (Really Simply Syndication) feeds help engineers, scientists, and other professionals to keep instantly up to date on advances in their fields. In short, the infrastructure of how information is stored and retrieved is changing rapidly.

Evaluation of Internet sites is an important consideration for anyone who uses this resource. An obvious, but often overlooked, mode of evaluation is to examine what person or entity produced the site. Many of the best sites for electrical and electronics engineers are produced by the same entities that produce the best print resources, such as professional societies, universities, publishers, and government agencies. An example of a valuable website for electrical engineers is that of the IEEE (http://www.ieee.org). The site presents a number of different options from the initial page, or home page, and it is not only attractive but easy to navigate as well. Information is grouped under intuitive headings, such as member services, and more than one path to a particular page within the site usually exists. The degree of editorial control is as important a consideration for evaluating a website as it is for a print source. As constructing a Web page is relatively easy, information in all gradations of quality and truthfulness shows up on the Internet. All information should be viewed with a critical and skeptical eye, which is especially true of Internet sites.

SUMMARY

Electrical engineering is a young discipline; the basic discoveries were only made in the 19th and early twentieth centuries. However, during the course of its existence, not only the infrastructure but also the scope of how information is communicated within the discipline has changed dramatically. When Michael Faraday first published his findings on the nature of electricity, they were available only to a tiny number of the scientifically curious. Now, scholarly and professional articles can be transmitted instantly to a potential audience of millions. The Internet, in particular, makes it possible for anyone who wishes to communicate scientific information to bypass the traditional gatekeepers of the dissemination of such knowledge. How this situation will be resolved is an open question. What is certain is that a revolution is occurring in the manner in which information is communicated. Many liken it to the revolution, both technical and social, that occurred in the aftermath of the invention of the printing press. Electrical engineers made many of the discoveries in electronics, computer hardware, and telecommunications, which led to this revolution. Increasingly, the discipline of electrical engineering will need to grapple with the changes in information retrieval and use that are occurring at an ever more rapid pace.

BIBLIOGRAPHY

1. McNeil, I. Basic Tools, Devices and Mechanisms.In *An Encyclopedia of the History of Technology*;McNeil, I., Ed.; Routledge: New York, 1990.
2. Parker, S. P., Ed.; *McGraw-Hill Dictionary of Engineering*. McGraw-Hill: New York, 1984.
3. Oppenheim, C. Patents and Patent Information.In *Information Sources in Engineering*, 3rd ed.;Mildren, K. W.;Hicks, P. J., Eds.; Bowker-Saur: London, 1996.
4. Knight-Ridder Information. *Complete Database Catalogue*, 1997. Knight-Ridder Information: Mountain View, CA, 1997.
5. Institution of Electrical Engineers. *A Classification Scheme for the INSPEC Database*. The Institution: London, 1995.

6. Holmberg, E. E., Ed.; *Gale Directory of Databases, Volume 1: Online Databases.* Gale Research: Detroit, MI, 1997.
7. University of Michigan. http://www.umich.edu/news/index.html?BG/google/index(accessed November 17, 2006).
8. Copyright Clearance Center, Inc. http://www.copyright.com/ (accessed November 16, 2006).
9. O'Leary, M. Google Scholar: What's In It for You? *Information Today* 2005, **22**,pp 35, 39.
10. Kinnaman, D. E. Sentimental Longings: Changing Education with Technology. *Technology Information* 1997, **17**,p 70.

MICHAEL CULBERTSON
Libraries Colorado State
University, Fort Collins, CO