Book Reviews

The Whole Internet User's Guide & Catalog. Sebastopol, CA: O'Reilly & Associates; 1992: xxiv + 376 pp. Price: $24.95. (ISBN 1-56592-025-2.)

Everyone with a computer seems to want to use the Internet these days—they've heard it mentioned in the newspapers, in the library, in classrooms, from their friends, and want to be a part of "it." The problem with the Internet is defining what "it" is, much less learning how to use it. The Internet is "the world's largest computing network," according to Krol. It is a loosely connected hierarchy of networks, with each network making connection agreements with the next network up the hierarchy, from local area networks (LANs) to campus or corporate networks to metropolitan area networks (MANs) to wide area networks (WANs), to national and international networks. While there are societies and coordinating boards, no one entity has final authority over "the net," making it quite difficult to learn what's out there. Nothing approaching an "internet catalog" yet exists, where one can search for people and information resources. As the resources proliferate, so do the tools to locate them—Gopher, Archie, Veronica, Jughead, Telnet, Hytelnet, Wide Area Information Servers (WAIS), World Wide Web, Campus Wide Information Servers (CWIS), lists of listservers, lists of online catalogs, etc. The novice is both excited about the possibilities and overwhelmed by the difficulties of knowing where to start.

Fortunately, some tutorial books, articles, and online workshops are now appearing. All of them seem to be written or organized by people who have learned the net through years of independent exploring and each has his or her own angle on how to present this mass of information to the novice in a coherent manner. Writing a book on "the whole internet" is a daunting task, not only because of the range of resources but because of the lack of consistency in interfaces to those resources. The Internet is a client-server world. Individual computer centers and computer operators have mounted servers that provide information. Generally speaking, the user's computer acts as a client that searches these servers, although a user may search a client, such as a gopher, which points to other gophers, and so on. The server may be on any type of operating system: UNIX and VMS currently are the most common, and many Macintosh, DOS, and other sites also exist on the net. Internet users may negotiate the net via a desktop system running any standard operating system, but capabilities vary greatly by whether the desktop computer is directly connected to the Internet ("a true Internet connection" in Krol's terms) or is making a dialup connection to a computer connected to the Internet. A university computer center may support Internet connections for several operating systems where the interface looks different to each, depending on whether it is a Macintosh, DOS, Windows, or UNIX system. Once the user connects to a remote system, each may see the same interface, if they are running the remote system in native mode, or may see an interface tailored to the local environment (e.g., a Macintosh direct manipulation interface client searching a UNIX server). So many combinations exist that it is difficult to explain the nature of the net in simple terms. Thus, to use the Internet one must learn basic concepts and functions such as electronic mail, file

transfer, and remote login; the next conceptual layer, tools and techniques to identify and locate people and resources; and how these functions and tools are implemented on your local system.

Krol has tackled this massive job and done some parts well and some parts not so well. The strength of Krol's book is its scope and detail. In 15 chapters plus appendices (xxiv + 376 pages), Krol covers all the basics and much beyond. He devotes a chapter each to the main functions of the Internet: electronic mail, remote login, and file transfer; a chapter on network news; and chapters on each of several major locator tools: finding software (Archie), finding individual people, Gopher ("tunneling through the Internet"), Wide Area Information Servers, and World Wide Web. The other chapters provide useful background information. "What is the Internet?" answers basic questions about funding and governance. "How the Internet Works" provides a clear explanation of packet switching, the IP and TCP protocols, and the domain name system; the latter is particularly useful in interpreting addresses and debugging access routes.

"What's allowed on the Internet" is a useful, if simplistic, discussion of law, politics, ethics, and security; it is enough to make the reader aware of these critical issues. Later chapters and appendices include "Other applications," which incorporates advanced features such as Rlogin and faxing over the Internet; "Dealing with Problems;" "Resources on the Internet," a catalog of basic sources; "Getting Connected," which briefly describes ways to get an account and differences in service levels; "International Network Connectivity," giving country codes; and "Acceptable Use," a reprint of the NSFNET Backbone Services Acceptable Use Policy.

The primary weakness of Krol's book is a mismatch with the audience. His stated audience was a "new graduate student in some non-technical discipline (i.e., not computer science or any form of engineering) who needed the Internet to do research" (p. xxi) and is computer-literate. The folksy tone of the book is appropriate to this audience but its UNIX-bias is not. Krol goes to great lengths to justify why basing all of the examples on UNIX should not be a problem and claims that "with minor variations, you'll find that they work the same way" on other operating systems. Computer literacy and UNIX-familiarity are synonymous *only* in the computer science and engineering disciplines; nontechnical users are highly unlikely to have UNIX interfaces on their desktops. A user must be more than "computer-literate" to know what "minor variations" might be required to translate from UNIX commands to other operating systems. If the UNIX examples consistently explained the actions or functions being executed, and gave Mac or DOS equivalents for at least some core functions, one would be more ready to forgive the UNIX-bias. However, he often gives a whole set of UNIX commands without even spelling out their meaning. In one of the first examples in the book he describes **ls** and **pwd** only as "command executed by remote system" and then refers to these in the explanation of the example as "standard UNIX commands" (pp. 46–47). This is insufficient information for the computer-literate novice to make the "minor variations" to another operating system. Some advice does not translate at all, such as recommending a mix of upper and lower case letters in a password—UNIX is the only standard operating system that is case-sensitive! Other advice is oriented toward system administrators rather than end users, such as suggesting that you examine the password file and list all the system tasks running if you suspect that your computer has been violated.

While the folksy tone of the book makes it readable, it tends to be verbose and lacks precise definitions of some critical concepts. The new Internet user seeking clarity on such concepts as client-server, OSI, or modem will not find them here. The glossary contains only 75 entries, several of which are cross-references to other entries. Some examples from the glossary:

client: A software application (q.v.) that works on your behalf to extract some service from a server somewhere on the network. Think of your telephone as a client and the telephone company as a server to get the idea.

OSI: Open Systems Interconnect; another set of network protocols.

modem: A piece of equipment that connects a computer to a data transmission line (typically a telephone line of some sort).

Krol discusses clients and servers in the context of telnet (pp. 47–48) by listing the actions that each entity takes, but nowhere in the book provides a concise definition of the form "a client is ... and a server is ..." that will allow the user to understand when they are running a client, or a server, or when they might seek better client software to run a particular server, etc. Client-server relations are so central to understanding the nature of the Internet that they deserve more explanation than they get here. Similarly, any definition of "modem" at least should explain that it stands for "modulate-demodulate" and translates analog telephone signals to digital computer signals and back; if nothing else, this much understanding is very helpful in interpreting those funny noises coming from the modem when it makes a phone call to another computer. The clear and detailed explanation of the TCP/IP protocols in the text and the lack of explanation of OSI protocols reaffirms the strong U.S. bias of the book as well (the author admits to this bias), as Europe is heavily OSI-based and is developing parallel standards to some of those discussed in this book.

Krol recommends that system administrators give this book to their users and tell them not to come back with Internet questions until they've read it. This reviewer concurs *only* if your community consists of sophisticated UNIX users who know much about computing and little about the Internet. If they indeed are nontechnical users and are users of Macintosh, DOS, or Windows, a better starting point would be *Crossing the Internet Threshold: An Instructional Handbook* (Tennant, R., Ober, J., Lipow, A. G.: Berkeley, CA: Library Solutions Press, 1993). While shorter and smaller in scope, the latter book offers a focused tutorial on information resources on the Internet, with exercises appropriate for users on several different operating systems and an extensive glossary with precise definitions. It is oriented toward library users, but suitable for a general computer-literate audience. *The Whole Internet* is suitable as the next addition to a basic library of Internet resources. The other essential book for a serious Internet user is *The Matrix: Computer Networks and Conferencing Systems Worldwide* (Quarterman, J. S.: Bedford, MA: Digital Press, 1990). *'The Matrix'* is the most complete reference source for architecture, protocols, international connectivity, networking history, and a variety of other topics; while the specific statistics are dated, it is so comprehensive (xxvii + 717 tightly edited pages) that it remains an indispensable resource on most aspects of networking. *'The Matrix'* also contains extensive bibliographies and indexes. *The Whole Internet* lacks a bibliography and contains only a minimal index and glossary. The most current sources about the Internet exist online, on the net itself—but you need to know the Internet well enough to be able to locate them, one of many catch 22s in the current state of networking. We are far from the day when we have a single, friendly interface that will provide a full-featured window to the Internet for all users. Until that day comes, we need ambitious books such as Krol's, despite their flaws.

Christine L. Borgman
*Graduate School of Library and
Information Science
GSLIS Building
University of California at Los Angeles
405 Hilgard Avenue
Los Angeles, CA 90024*

Operations Research for Libraries and Information Agencies. Donald H. Kraft and Bert R. Royce. San Diego, CA: Academic Press; 1991: 193 pp. Price: $57.50. (ISBN 0-12-424520X.)

The philosopher, Robert Nozick, in his latest book (Nozick, 1993), mentions a peculiar anomaly.

When the Encylopedia Britannica published its (second) edition of "Great Books of the Western World," this occasioned some public controversy over the representation ... of women and minorities ... What received no comment however, was that many of the greatest intellectual works of the twentieth century were omitted, presumably because they were too technical for the intelligent generally educated reader.

Information studies faces a similar problem. Some of the most important work done in the last 20 years is essentially technical in nature. Yet most programs in Information Studies require little or no technical preparation for admission to graduate studies. The result is that we are teaching with one hand tied behind us.

I have been very slow in reviewing this book, as I sought to understand where it fits into the educational and research arena of Information Studies. The book is a good introduction to a range of topics, including systems analysis, decision theory, graph theory, queuing theory, stochastic processes, simulation, and modeling. It begins with an overview and justification of operations research in the library field, and closes with a discussion of how this is to be implemented in the aid of library management. Chapters 1–5 could be used, by a teacher who is comfortable with this sort of material, as part of a broader course in library management, or in an advanced seminar.

In my own teaching I have found that the spreadsheet has given us a powerful tool to overcome both fear and ignorance of algebra. When a calculation can be set up in spreadsheet form, students master the notion of a "variable, x," as they change the contents of the cell corresponding to x, and see the other parts of the spreadsheet change accordingly. Hence, many concepts which we have previously portrayed only with the use of algebra can, in fact, be brought into the classroom using the power of the computer.

Thus the work by Kraft and Boyce gives us a powerful tool to wield in the hand now tied behind us. It shows a path down which those of us who are comfortable with this essential and technical component of information studies can direct the preparation and education of graduate students. In a nutshell, if you are the sort of person who can read the equation in Figure 1 of this review, and follow its meaning as the general scholar will follow a complex verbal argument, this text will be a great help. It can serve as a the focus for an advanced seminar for technically oriented students in an MLS program, and for doctoral students.

Much of the work is a survey of topics which can be studied in greater detail elsewhere, and extensive references are given. Some

$$p\left(g^{*}_{j,t,q} = b \mid g^{*}_{j,t+1,q-1} = a\right)$$

$$= p\left(g^{*}_{j,t+1,q-1} + \alpha_{j,t,q}n_{t,q} = b \mid g^{*}_{j,t+1,q-1} = a\right)$$

$$= p\left(\alpha_{j,t,q}n_{t,q} = b - a \mid g^{*}_{j,t+1,q-1} = a\right)$$

for all j, t, and q.

FIG. 1. A typical "challenging" expression from *Operations Research for Libraries and Information agencies*. $g^{*}_{j,t,g}$ represents the number of relevant articles appearing in journal j, in time period q, which are still relevant at the later time period $q + t$. $n_{t,q}$ is the total number of articles published in q and still relevant at $q + t$. $\alpha_{j,t,g}$ represents the fraction of those that are in journal j.

of the material, especially chapter 7, is essentially original work by the authors. Thus the book is at the same time a textbook and a research monograph. The book is generally free from errors.

If I find any fault, it is with the extensive listing of software for operations research, without annotations as to source, scope, scale, or price. This is a rapidly moving field, and the educator would do better to peruse the pages of *OR/MS Today* (1994) for both advertisements and reviews of educational software.

In addition, substantial constrained nonlinear optimization capability is available in Borland's Quattro-Pro Version 4.0. A final quibble: The formulae on p. 108, while true, do not well illustrate the matrix multiplication of the first order Markov model.

In summary, I think this book should be on the shelves of all those who can read it, and that it should be used to increase the number of such people in the next generation of Information Studies/Science scholars and researchers.

Paul B. Kantor
SCILS
Rutgers University
Huntington Street
New Brunswick, NJ 08903

References

Nozick, R. (1993). *The nature of rationality*. Princeton, NJ: Princeton University Press.
OR/MS Today. Baltimore, MD: Operations Research Society of America and The Institute of Management Sciences (ISSN 0030-3658).