TER Volume 4, Issue 8, September 1997

Telecommunications Electronic Reviews (TER) is a publication of the Library and Information Technology Association.

Contents:

ter issues (/lita/publications/archive/ter)

REVIEW OF: Clinton Wong. Web Client Programming with Perl. Sebastopol, CA: O'Reilly & Associates, 1997.

by Mark Cyzyk

The purpose of this book is to introduce UNIX-based Webmasters to simple Web client programming using Perl (Practical Extraction and Report Language). It can, however, be profitably read by anyone interested in how Web clients work.

The first third of the volume concentrates on what Web clients are and how they do what they do. The second chapter, "Demystifying the Browser," examines the anatomy of a Web client and the basic structure of a client/server transaction. The various HTTP (Hypertext Transfer Protocol) headers that are passed from client to server and back again are illustrated. To illustrate the text-based nature of HTTP, Wong offers an interesting example of a manual HTTP transaction. He shows how one can simply Telnet into a Web server and manually issue HTTP commands resulting in output from the server. This powerful example of a simple HTTP transaction stays firmly in the reader's mind during the rest of the book.

The third chapter, "Client Request Methods," offers a relatively in-depth discussion of HTTP itself. Although there are fully ten methods that a client can use to make a request of a server, Wong spends time only on the seven most important: GET, HEAD, POST, PUT, DELETE, TRACE, and OPTIONS. These are the most useful request methods for the budding client programmer to understand. All CGI (Common Gateway Interface) programmers are familiar with the GET and POST methods, but Wong's discussion of the other methods is fascinating to those like myself who, when confronted with the automatic upload features of recent versions of Netscape and other Web-based applications, wonder "How did they do that?" (Hint: they used the PUT method.) The chapter ends with a solid discussion of server response codes, e.g., 403 Forbidden, and what they mean, as well as a comprehensive list and select discussion of HTTP headers.

For the client programmer who revels in the bits and bytes, chapter four, "The Socket Library," covers network socket programming over a TCP/IP (Transmission Control Protocol/Internet Protocol) network, specifically in a UNIX environment. Though important for a comprehensive understanding of the client/server process over TCP/IP, as Wong himself states, a deep understanding of socket programming is not needed to create Web-based client programs. The reason for this is that programmers can avail themselves of the Library of World Wide Web (WWW) modules for Perl (LWP) that are widely available from most Perl archives.

Like any Perl library, the LWP modules exist to greatly simplify the programming process. The heart of the book is to be found in chapters five and six, "The LWP Library" and "Example LWP Programs," where the LWP Library is discussed and its uses are illustrated. LWP is a library of modules for Perl 5 that allows the programmer to accomplish complex tasks (such as socket programming or Uniform Resource Locator parsing) in a simple and elegant manner.

The library as a whole is comprised of eight discrete modules of which Wong addresses only the four most useful: the LWP, HTML (Hypertext Markup Language), HTTP, and URI (Uniform Resource Identification) modules. Essentially, the LWP module handles socket communications between client and server; the HTML module handles the parsing of HTML documents; the HTTP module handles HTTP requests and responses; and the URI module handles the escaping of URIs and the relative-to-absolute translation of URLs. Within each module are to be found several classes (LWP is object-oriented); Wong discusses the most important for each of the four modules.

Wong offers three types of example Web clients--simple clients, periodic clients, and recursive clients--to illustrate the many uses of the LWP library. Wong's simple client calls a URL with a custom User Agent header designating the name of the client program, thereby identifying itself to the server. His periodic client--more complex than the simple client--periodically connects to the Federal Express Web site, uses a tracking number to query its database, and waits to see if a package has been delivered. The program runs

until it receives notification of delivery. Such a periodic client could form the basis of any Web-based function that needs to be performed regularly, for example, a stock ticker checker. Wong's recursive client-- by far the most complex--checks all the HTML pages at a specified Web site for malfunctioning hyperlinks. A list of these links is then printed out.

For those running X Windows, the final chapter, "Graphical Examples with Perl/Tk," illustrates how one can program a graphical user interface (GUI) for Perl text-based programs. In this manner, the programmer can actually program a custom GUI Web browser.

As are most of the O'Reilly publications, this book is well- prepared, useful, and a pleasure to read, using the chatty style that O'Reilly authors seem to prefer. Examples throughout are relevant and clearly presented. The book as a whole proceeds in a satisfying manner, its pace easing the reader into ever deeper programming concepts and examples. For the Webmaster hoping to create custom, text-based Web clients for use in Web site administration, or for those just interested in learning more about how the hypertext transfer protocol functions, this book is recommended.

Mark Cyzyk (mcyzyk@towson.edu (mailto:mcyzyk@towson.edu)) is Head of Information Technology at Albert S. Cook Library, Towson State University, in Maryland.

table of contents ter issues (/lita/publications/archive/ter)

REVIEW OF: Daniel J. Barrett. Net Research: Finding Information Online. Sebastopol, CA: O'Reilly, 1997.

by Mary Lynn Rice-Lively

"Dan Barrett attempted his first search on the Internet in 1985, when he typed 'ftp *.*' and received an indecipherable error message." (p. vii)

Who among us has not experienced the indecipherable when searching the Internet? Net Research, a "reader friendly" guide to Internet research strategies, offers proof that techniques do exist to perform productive Internet searches. Unlike print publications that merely list ephemeral Uniform Resource Locators (URLs) of Web pages, Net Research, as the author notes, "is built to last." (p. xii) Readers exhibiting all levels of Internet expertise and users of commercial Internet Service Providers (ISPs) will be rewarded with search strategies and techniques that will remain useful even as the Internet evolves.

How many of us can articulately answer the question, "What is a search engine?" Although Barrett explains the difference between passive (Web sites are registered with the search site) and active (using crawlers and robots to retrieve information) search sites, he does not clearly define the tool. A search engine is "the popularly recognized group of tools that have been developed recently (in the past few years) to index and retrieve information stored on the Internet." [1] Others might include in this definition the programming that

supports searches, the database of information, and the interfaces that permit users to enter their searches. With this in mind, Barrett's guide is more of a "how to" search guide than a technical or even a location resource.

Experienced technical trainers and instructors might agree that order and clarity contribute to understandable and approachable instruction. Net Research, divided into four sections, supports this tenet by providing in the first four chapters a conceptual framework for the Internet searcher. This includes an exploration of the basic tenets of searching and how the Internet is organized.

Tempting even the most skeptical of Internet users, chapter one begins the book with an optimistic and practical "Internet Searcher's Rules for the Road." Among Barrett's six rules for searching success are tips such as "Don't assume failure too quickly" and "Think about your route." Even the unseasoned searcher knows that, with regard to the Internet, encouragement must be followed by admonitions. Barrett's list of admonitions are no exception.

Barrett's instructional experience shines through in his use in each chapter of clear objectives at the beginning. Each chapter (beginning with chapter two) concludes with a quiz to guide the reader in a "hands-on" practice of the techniques explored in the previous discussion.

Chapter two, "Internet Basics," and chapter three, "Views of the Internet," may provide redundant information for the seasoned Internet user. The summaries prepare all levels of Internet users to begin the explorations of searching with a succinct, clear, and visually understandable explanation of protocols encountered while using the Web (e.g., HyperText Transfer Protocol (HTTP) documents, Gopher, FTP (File Transfer Protocol), and Usenet information.) Even the experienced Internet user should read chapter four's tips on "Choosing an Effective Starting Point" for the reminders that all search engines: are not alike, do not search the same information, are not always the best place to begin a search, and do not provide the same search options.

Many among us who are online instructors and trainers begin preparation for a class by scanning the Web for examples of tutorials and class syllabi. For example, Ross Tyner's online tutorial, prepared for the British Columbia's Educators Workshop on Educational Technology, "Sink or Swim: Internet Search Tools & Techniques" provides an excellent Web-based guide to teaching and learning to search online (http://www.sci.ouc.bc.ca/libr/connect96/search.htm (http://www.sci.ouc.bc.ca/libr/connect96/search.htm)). While the guide is complete with strategies using Boolean logic and simple queries, it is an online, not a print-based guide.

Another online, high-powered search guide is Tracy Mark's Windweaver (http://www.windweaver.com/ (http://www.windweaver.com/)) that provides guides for new users as well as rankings of the search engines. Despite the excellent examples of tutorials, search tips and strategies available on the Web, Barrett's book serves as a physical (as opposed to digital) text for individuals working to improve their search results and efficiency. Other useful guides to searching include Internic's Search Tutorial (http://rs.internic.net/nic-support/15min/ (http://rs.internic.net/nic-support/15min/)) and Richard Smith's Web Cheat Sheet (http://www.colosys.net/search/ (http://www.colosys.net/search/)), which includes convenient access to the help guides and forms interfaces of seven of his favorite search engines.

Net Research's methodical dissection (non-technical readers need not be squeamish here) of how search engines work is conversational, guiding the reader through a thoughtful analysis of individual search preferences (e.g., simple or advanced queries, natural language or Boolean logic.) For the newcomer to precise and thoughtfully constructed search logic, a careful reading of Net Research will, no doubt, contribute to improved CD-ROM as well as Internet search strategies.

The focus of chapters six through nine is finding specific kinds of information: people, places, topical discussion groups and newsgroups, and software. The concluding chapters advise of the strengths and weaknesses of bookmarks for "Finding Information Again."

An additional feature is the guide to options for constructing private Web pages with instructions for several commercial ISPs (e.g., America Online, CompuServe, Microsoft Network, and Prodigy). Coverage is also provided for those who have full UNIX access and want to create their own Web pages, including how to inform "passive" search sites that a page is there. Most of this information (e.g., answering questions in a discussion group, making FAQs (Frequently Asked Questions), running a mailing list) duplicates other more general guides to the Internet.

Net Research: Finding Information Online continues the O'Reilly publishing tradition of guiding readers with precise and usable Internet instruction and analysis. On one hand, the analogy still holds true that Internet information is as organized as a library with all of its resources stacked in a giant pile. On the other hand, a searcher who takes the time to read Barrett's Net Research can make productive use of today's increasingly sophisticated search tools.

As an online bonus you can access Barrett's "Great Places to Start a Search" at O'Reilly's Net Research Web site (http://www.ora.com/catalog/netresearch/ (http://www.ora.com/catalog/netresearch/)). While greatly improved since Barrett's first FTP search in 1985, Internet searching continues to be part art and part science. Be forewarned when firing up your computer and Web browser to deliver your search strategies to bots, worms, crawlers, and the like, not to abandon your intuition and common sense.

Notes:

[1] Brandt, D. Scott. (1997, January). What Flavor is Your Search Engine? Computers In Libraries, 17(1), 47-50.

Dr. Mary Lynn Rice-Lively (MARYLYNN@MAIL.UTEXAS.EDU (mailto:MARYLYNN@MAIL.UTEXAS.EDU)) is the Coordinator of Information Technology for the Graduate School of Library and Information Science at The University of Texas at Austin.

REVIEW OF: Karen G. Schneider. The Internet Access Cookbook: A Librarian's Commonsense Guide to Low-Cost Connections. New York: Neal-Schuman Publishers, 1996.

by David Mattison

A well-known name among certain librarian-oriented mailing lists, Karen Schneider is also "a genuine MLIS-librarian ... and a freelance Internet trainer." (p. vii) Her book began as a handout at various training sessions and is intended for the computer novice. The book is geared to individual librarians planning to

obtain dial-up access to the Internet.

While Schneider is relatively thorough in covering the basics of acquiring a computer, installing a modem, and comparing the two different types of dial-up access, she barely discusses the basics of what to do once you're online. For example, in a sidebar on using FTP (File Transfer Protocol) to download software, she doesn't mention the "ls" command needed to list the files in directory in order to make sure the file you want is there. A brief overview of the few UNIX commands necessary to adequately use a shell account (primarily found on UNIX systems) would also have made the book more self-contained.

Schneider covers both DOS/Windows PC systems and Apple Macintosh computers. While she favors the latter for ease of general set-up, she admits that configuring a Mac for Internet access can be tricky. Windows95 was just becoming available as she wrote the book. My own experience with Win95 and a Mac PowerPC would cause me to favor the former for establishing a dial-up Internet connection using PPP (Point-to-Point Protocol). TCP/IP (Transmission Control Protocol/Internet Protocol) and PPP are fully integrated into Win95 and, assuming Win95 was installed correctly, configuring for Internet dial-up access is extremely simple.

The book is divided into four parts covering:

1. overview of Internet and types of Internet accounts,
2. computer hardware and software,
3. requirements for setting up a shell account and a SLIP (Serial Line Internet Protocol) or PPP account/connection, and
4. connectivity costs and print/online resources.

Four appendices encompass commercial Internet service providers, Free-Nets (community computer networks), mail-order computer vendors, and a glossary. The index is just over three pages and not totally consistent in content coverage. There are many illustrations, including screen shots, but perhaps not enough for the novice. Most of the photographs are of Mac equipment and most are printed as negative rather than positive images!

Part four, which provides some important economic models for low-cost connectivity, could have been eliminated by combining it with chapter two where the different types of Internet accounts are discussed. While there is some repetition among the different parts, Schneider uses these recurring topics as reinforcement and as a form of hypertext for those who tend to jump around (something she encourages).

Schneider's advice is, on the whole, sound, based on her experience and solid reference works. She writes in a breezy style, but the cookbook analogy wears thin very quickly. My overall impression is that this book is too anecdotal and not sufficiently technical.

The chapter on purchasing computers is particularly good because of the pro-and-con approach she uses for each type of purchase source. Schneider personally prefers mail-order shopping and suggests studying ads and phoning toll-free numbers; she notes, "That is, quite seriously, how I taught myself a lot of what I know." (p. 74) Schneider finds the mail-order approach especially suitable for locating new Mac software and peripherals.

The overall Mac slant and the relatively slow rate of software change for that platform mean the currency of this book for Mac users is still quite high. For PC users, however, the introduction and meteoric impact of Windows95, particularly with respect to Web browsers, suggest that this book's value is about as high as a

286 PC. Parts of it still work well, but buying a computer with Win95 with integrated TCP/IP and PPP provides simpler Internet configuration and the ability to keep up with rapid software changes.

One personally pleasing aspect to this book is the author's inclusion of Free-Nets as a significant connectivity option. Libraries and Free-Nets go hand in hand. Sadly, however, the National Public Telecomputing Network that oversaw the development of the Free-Net community filed for bankruptcy early in 1997. Nevertheless, the overall strength of the community computer network movement is such that most seem capable of sustaining themselves amidst commercialization pressures. The partnerships between community-based computer networks and libraries are part of that sustenance.

Even though this book was written by a librarian and Internet evangelist for her colleagues, the information that is presented is not, unfortunately, of sufficient distinction to make this a purchase recommendation. More inclusive, less expensive print resources such as The Complete Idiot's Guide to the Internet, The Internet for Dummies, or The Internet Starter Kit books provide the necessary basics to get connected. Once connected, it is relatively simple to find the economic data an information professional might need to convince a hardened, bottom-line manager that Net access is a necessity these days, not a luxury. Take one of Schneider's approaches: join a popular mailing list like Public-Access Computer Systems for Libraries (PACS-L) and ask.

David Mattison (David.Mattison@gems3.gov.bc.ca (mailto:David.Mattison@gems3.gov.bc.ca)) is Reference Archivist at the British Columbia Archives in Victoria, Canada.

table of contents   ter issues (/lita/publications/archive/ter)

REVIEW OF: Allen C. Benson. The Complete Internet Companion for Librarians. New York: Neal-Schuman Publishers, 1995.

by Leroy D. Smith

Allen Benson's Complete Internet Companion for Librarians was an excellent introduction to the Internet when it was published nearly three years ago. Benson, the director of the Saline County Library in Arkansas, set out to help "[l]ibrarians who want to integrate Internet services into their reference practice ... for locating specific resources listed by subject or for learning more about resource discovery tools," and for the most part he was successful. (p. xxi) His "how to" chapters are filled with examples, step-by-step instructions, sample activities to try, and pointers toward further study. His descriptions of FTP (File Transfer Protocol), Telnet, Gopher, WAIS (Wide Area Information Server), and Archie are as clear as any I've previously read.

Unfortunately, it is difficult for any book based on information four or five years old to hold up well with regard to a subject as rapidly changing as the Internet, where the life of specific versions of software or particular resource collections is usually measured in weeks and months, not years. Because of this, most

of the information in Benson's book now falls into one of three categories: 1) it is out-of-date and inaccurate, or 2) it is valid theoretically but specific references and examples are dated and in many cases no longer accessible, or 3) it is still accurate, but improvements and new software have rendered it obsolete.

User-friendly interfaces like Netscape or Internet Explorer, for example, have made much of Benson's book unnecessary. It is still possible to laboriously FTP files in the manner Benson describes, but the same results are attainable so much more easily through a browser like Netscape that it is difficult to imagine anyone using the old method. Many of the Gopher sites Benson mentions are no longer in service, and those that still exist seem to have stopped adding information about the time Benson's book was published (many refer users to their Web sites for the most accurate information). NASA's Spacelink (one of Benson's examples of the usefulness of Telnet) can no longer be reached by Telnet; it, too, is only available now through the World Wide Web.

A few chapters of Benson's book are still pertinent. His explanations of the uses of email, listservs, and newsgroups, for example, are useful. But even in this area the specifics are outdated: at the time the book was published many of the current versions of mailing list software (e.g., majordomo, listproc, mailserv, etc.) were not in use, most listservs were still using BITNET addresses, and his list of library-related listservs is only partially complete. In the appendices, Benson provides nearly 40 pages listing specific Internet resources, arranged by subject, but unfortunately most of these are Gopher and FTP sites which have been bypassed or incorporated into similar sites on the World Wide Web.

The bottom line is that today any book about the Internet which does not mention Netscape, which does not discuss how to locate information on the Internet using search engines like Alta Vista or hierarchical subject directories like Yahoo, which does not grapple with the problem of evaluating resources obtained through the World Wide Web, such a book cannot be said to be "complete."

Benson had the right idea. "The Internet has created new possibilities for libraries, from the smallest to the largest, enabling them to participate in global resource sharing. The Internet has successfully provided greater access to more information resources and services ...." (p. xx) Many librarians need some specific help and guidance in making use of this resource. They need help locating resources, they need guidance in the appropriate use of the Internet in reference service, they need help organizing training for their patrons, they need help in terms of copyright issues and free-access issues, they need help in grappling with the questions of print vs. electronic journals and cataloging Internet resources, they need help using the Internet as an internal resource for the operation of their libraries--these are all things Benson tries to do with his book.

However, above all, they need the most current information available. If ever there was a book in need of a second edition, worthy of being revised and updated, this is it. But until it is updated, librarians would be wise to think twice before purchasing it.

Leroy D. Smith (LSmith@cc.colorado.edu (mailto:LSmith@cc.colorado.edu)) is Reference Librarian at the Tutt Library of The Colorado College.

REVIEW OF: Xia Li and Nancy B. Crane. Electronic Styles: A Handbook for Citing Electronic Information. Medford, NJ: Information Today, Inc., 1996.

by Kristin Vogel

Containing 14 chapters and over 500 examples, this book is a revised edition of the 1993 book of the same title. The preface states that the revision grew out of the introduction of new technologies, requests for MLA-style citations, and the authors' desire to more fully develop their ideas.

This revised edition contains guides to both the American Psychological Association and the Modern Language Association citation styles. In addition to these two styles the authors also incorporate recommendations of The Bluebook: A Uniform System of Citation for Legal References. The authors use the language "APA embellished" and "MLA embellished" citations to refer to their styles in order to indicate that they have made some recommendations that do not follow the exact guides of the original styles.

Chapters one and eight provide an overview of each style, modifications from the Associations' recommendations and from the earlier edition, and principles for citing materials in the respective styles. In these chapters, the authors work through each element of the citation providing examples and stating rules on spelling, punctuation, and capitalization. Part one (chapters one through seven) covers the APA embellished style, while part two (chapters eight through fourteen) covers the MLA embellished style.

Chapters two and nine cover items such as individual works, books, monographs, and full-length works found in full-text databases. Full-text databases include CD-ROM, commercial online databases, email, File Transfer Protocol, Gopher, Hypertext Transfer Protocol, Telnet, USENET, and Wide Area Information Servers. In these chapters the authors begin use of their convention of using an alphanumeric identifier for each of the 276 examples applied to these styles. The identifier uses A to refer to an APA embellished example and M to refer to an MLA embellished example. In this and each subsequent chapter the alphanumeric identifier provides the reader with the capability of comparing the two examples.

Chapters three and ten provide information on citing periodicals, delineated into journal, magazine, newsletter, and newspaper articles. Chapters four and eleven outline discussion lists, USENET newsgroups, and personal mail. Chapters five and twelve address United States government documents and legal sources as well as international documents. Chapters six and thirteen provide additional examples of less common items that may be cited, including full CD-ROM databases, entire World Wide Web Homepages, audiovisual materials, wire service reports, and abstracts.

Chapters seven and fourteen illustrate rules of documentation for the respective styles. The chapters show single, multiple, and corporate authorship along with no author and multi-work citations. This section also addresses personal communication through email or online postings. At the end of each part, the authors have included reference lists corresponding to the respective style. Following the works cited list of part two is the 14-page index.

This revised edition is a marked improvement on the first edition. The sections on MLA style are a strong addition to resources available about citing electronic resources. This book provides the depth and comprehensiveness needed for accurate citing of resources in both APA and MLA styles.

The inclusion of chapters one and eight is definitely a strength of this text. These chapters demonstrate and give rationale for any differences from the official association recommendations of APA and MLA styles. For an individual using this style manual in preparation of a manuscript, this section is very useful in cases

where questions arise about the use of particular citation elements and rules of spelling, punctuation or capitalization. The authors are very clear in stating their case for the full construction of electronic citations.

The book is not an easy resource for beginners in research and requires familiarity with electronic publishing. To use this book, an individual must first know the definition of the elements of an electronic address and where to find them in an online resource. In particular, access date is not necessarily a standard item recorded by students while researching. Students must also know the difference between "no date" and "no access date."

Thorough knowledge of the many varieties of electronic resources is also helpful in becoming familiar with the authors' categories. Students, or those not thoroughly familiar with electronic resources, may be confused initially. Technically, all of the sites accessible with HTTP could be considered World Wide Web pages. However, the authors use the term World Wide Web pages to refer to a specific type of page on the Web. HTTP is used to refer to many other sites on the Web such as electronic journals or newsletter articles.

With an awareness of the prerequisites for successful use of the text, I strongly recommend this book as a resource for any reference collection. The book is very well organized. There are many examples included to address the vast majority of resources accessed electronically. An extensive index is also included. This book is a must for those providing instruction in the use of electronic resources.

The first edition of Electronic Styles was a pioneering work in electronic citations. This edition has taken strides in the direction of even greater usability. As citations of electronic resources become commonplace in research, this book stands out as a standard and necessary reference resource.

Kristin Vogel (kvogel@titan.iwu.edu (mailto:kvogel@titan.iwu.edu)) is Access Librarian and Assistant Professor at Illinois Wesleyan University.

About TER