

program presented chronologically from beginning to end. Despite these structural issues, the advice offered in each chapter is relevant. The author's discussion of strategic planning, for instance, addresses issues such as user needs assessment, internal resource audits, mission statements, and vision documents. It also provides example documents, such as a user needs worksheet, internal needs audit worksheet, and sample mission statement, which are helpful to the project planner. These examples illustrate key points in the text, support and elucidate the author's objectives, and are consistent throughout the book.

In the "Technical Overview" chapter, Clobridge addresses the critical area of budgeting. She addresses ways to curtail costs, such as using open source software, sharing and purchasing refurbished equipment, outsourcing, and batch processing. Staffing is another significant financial consideration in building and maintaining a digital repository program. Clobridge emphasizes the importance of interdepartmental teamwork across the institution. She is correct in her assessment regarding the critical need for committed personnel, not just librarians, library support staff, and student workers, but also library administrators, faculty, and information technologists. Her text stresses the substantial time commitment required institution-wide for digital repository programs to succeed. Therefore, although the author's intent is to write a guide for all academic institutions of all sizes, it is not as useful for small libraries with few personnel. For medium-to-large colleges and universities with adequate staff resources, this handbook functions as an appropriate and excellent guide.

Clobridge's account of metadata in chapter 5 offers a concise overview of the different types. She briefly describes Dublin Core (DC) and the Visual Resources Association (VRA) Core, which are highly respected and pervasively used standards. DC is

flexible and works for different types of collections, while VRA Core is used with art collections. She briefly lists others, including Text Encoding Initiative (TEI), Content Standard for Digital Geospatial Metadata (CSDGM), and Encoded Archival Description (EAD). The author provides an overarching description and gives references to specific instructions for each standard. She offers advice on selecting the appropriate markup language, ways to streamline production, metadata cleanup, and enhancements. She advocates progress over perfection, and harvesting metadata when possible. In this chapter, the author focuses more on the process of applying metadata than to its technical details. Definitions pertaining to metadata work are presented and are helpful to those new to markup languages.

Another technical discussion appropriately included in the handbook is the sustainability of digital data. Librarians must be concerned with digital preservation as more information is converted to or born digital. Clobridge speaks to these issues with good advice about preservation planning, data backup, addressing obsolescence, disaster planning, managing electronic records, licensing and legal issues, and supporting collections financially over the long term.

Clobridge's work serves as an effective starting point for planning, preparing, implementing, and maintaining digital collections in an institutional repository program. The book imparts the necessary knowledge to plan and begin such projects, and it directs readers to useful sources for detailed manuals and instructions. —Shannon Fox (*sfox@austincollege.edu*), *Austin College, Sherman, Texas*

***Introduction to Modern Information Retrieval.*** 3rd ed. By G. G. Chowdhury. London: Facet, 2010. 508p. \$73 softbound (ISBN 978-1-8560-4694-7).

Information retrieval is a large and evolving topic. G. G. Chowdhury's

*Introduction to Modern Information Retrieval* was first published in 1999 and now enters its third edition. The author approaches the topic from a library perspective, but endeavors to place library practice in the context of broader concerns in information retrieval and resource discovery generally. The book covers topics such as metadata formats, subject analysis and vocabulary control, abstracting and indexing, database technology, file organization, user requirements and interface design, natural language processing, digital libraries, web search engines and search engine optimization, evaluation of retrieval effectiveness, and experimental research.

One hardly need point out that many changes in information retrieval have taken place since the first edition of this book was published. This new edition includes new or expanded sections on developments in metadata standards, such as the Functional Requirements of Bibliographic Records (FRBR), automatic indexing, information retrieval on the web, XML markup and retrieval, and citation analysis.

The book draws heavily on the published literature. Each chapter has an introduction followed by an outline of important ideas in the area under discussion, often including summaries of significant published works. Each chapter concludes with a discussion and a lengthy list of references. This structure sometimes gives the book the feeling of an extended literature review rather than an independently argued viewpoint, but readers using this publication as a textbook will appreciate key contributions in the field being brought together and having both their significance and content succinctly explained. The discussion sections, seldom consisting of more than a single paragraph, summarize main themes covered in the chapter. The volume is furnished with a subject index.

The book's wide scope is its greatest strength. The author treats

information retrieval not as an abstract problem of conceptual design but a practical one requiring an understanding of user needs as well as general design principles and specific technologies. The book serves its readers particularly well as a primer on the basic technical aspects of computerized information retrieval. The chapter on automatic indexing and file organization, for example, goes into considerable detail about alternative methods of term weighting and about how inverted files and alternative text retrieval structures work. The chapter on searching and retrieval includes not only information about query formulation and search strategies but also an explanation of vector processing. All of these topics are essential background for anyone working in this field.

I would have liked to see metadata issues analyzed in greater depth. Although the book covers traditional tools such as classification schemes and subject headings at some length, it does not fully articulate the rationale for recent developments, including FRBR and the proposed new cataloging code, Resource Description and Access (RDA). Missing is a deeper interest in exploring how metadata schemes and interface design interact in an effort to solve retrieval problems. For example, the chapter on cataloging and metadata contains screen shots of a faceted catalog display but relatively little analysis of how such displays

utilize MARC, metadata originally designed for a different environment. The chapter on multimedia information retrieval outlines some high-level issues pertaining to music retrieval, but a closer study of something like the Indiana Variations<sup>2</sup> project, briefly mentioned in the text, would have given readers a stronger sense of the specific challenges in this domain.

The question for any revision of an established textbook is how well it covers recent developments. Here I think the book meets with only qualified success. Recent years have seen a trend to decouple the discovery layer from the library management systems used to maintain the underlying metadata, a trend that comes as part of the effort to broaden the range of content that the discovery interface can encompass. This significant development is not specifically discussed in the book. The section in question instead focuses mainly on earlier approaches to the problem based on broadcast searching. Link resolution, an important technology in this type of environment, receives only a passing mention. These gaps are surprising given that the author devotes an entire chapter (chapter 22) to digital libraries, the types of aggregations they embody, and the issues heterogeneous aggregations pose for information retrieval. Proponents of the Semantic Web will be disappointed to find that their topic is not treated.

Other topics of current interest in web retrieval, such as search engine optimization and metadata harvesting, are acknowledged, but receive only limited attention.

Some of the text should have received more careful proofreading. For example, the otherwise very valuable chapter on retrieval evaluation experiments cites the 2002 Text Retrieval Conference (TREC) on page 312 as being the most recent, but then goes on to mention the 2009 workshop on the next page. (The table of TREC tracks on page 314 has been updated as far as 2002, but no further, perhaps a reflection of the 2004 publication date of the book's second edition.) The book includes a handful of screen shots of end user interfaces, but at least in my review copy, the images (for example on pages 54–56) are often so blurred as to be almost unreadable. I hope this problem has been corrected in production copies.

Part of the difficulty of keeping up with developments in resource discovery is that the scope of the problem keeps changing. Chowdhury's book nevertheless deals expertly enough with a broad range of the basic issues to remain a valuable resource for students and practitioners. This new edition will extend the book's useful life.—*Chew Chiat Naun (chewx006@umn.edu), University of Minnesota, Minneapolis*