
by Kwan-Yau Lam

As its title clearly indicates, this book is about developing powerful and robust Web applications by integrating the strengths of XML (Extensible Markup Language) and SQL (Structured Query Language—a language popularly used to query, that is, to extract data from, databases). Essentially, the author argues that XML is a wonderful tool for structuring data, and relational databases are good at storing and relating data, hence integrating the two could result in applications with immense power. In this book, the author has also shared with his readers many of the ideas and insights that he gained from his work as a technology consultant and a content management specialist.

This book is intended for software developers who work on small- or Medium-scale projects and who often have to face resource limitations of one kind or another. As noted by the author, familiarity with concepts of relational databases and markup languages, particularly HTML (HyperText Markup Language), would certainly make it easier for the reader to grasp the contents of this book. In addition, this reviewer believes that some knowledge and experience in computer programming would help one follow most of the XML and SQL source codes found throughout the book.
Note that this book is not written as a comprehensive user's guide. Any reader who expects to find lists of XML or SQL commands with detailed explanation in this book would surely be disappointed. In fact, the author has opted to take a rather conceptual approach in discussing his ideas, concerns, and issues related to Web application development using markup languages and relational databases.

The organization of the book is quite clear and straightforward. The book contains ten chapters, which follow a logical sequence and which, according to the author, correspond roughly to different stages of designing and developing Web applications. As a matter of fact, the ten chapters can be further divided into three parts: (1) introduction and history, (2) XML application development life cycle, and (3) standards, frameworks, and examples for building applications.

The first part, consisting of chapters 1 and 2, asks the question "Why XML?" as an introduction to the potential of XML and presents a brief historical account of XML and SQL. The basics of XML, its evolution from SGML (Standard Generalized Markup Language), and the advantages of XML over HTML are some important topics covered here. Overall, this part provides some good basic background information and helps motivate the reader to go on reading the latter chapters.

The second part, consisting of chapters 3 through 6, is perhaps the most important part and the core of the entire book. This is where the main theme of the book--that the marriage of XML and relational databases would give immense power to applications--is argued and explained. Topics covered in this part follow a logical order, which also reveals the different stages in the life cycle of XML application development--first from project definition and management, then to data modeling, then to XML design, and finally to database schema design.

As mentioned earlier, the author has chosen to take a conceptual approach in this book. This conceptual approach is most apparent in part 2. Using two fictitious application development scenarios (e-mail and a movie review database) as examples, the author focuses on important concepts underlying each of the four development stages outlined above. Note, however, that not all of the concepts discussed are necessarily technical in nature. Some are more pragmatic and logistic than technical. For example, in discussing the first stage of project definition and management, the author has put much emphasis on the importance of thorough gathering, analysis, and documentation of functional requirements from the perspectives of actual end users and administrators.

The author has also stressed the importance of building an abstract, application-independent data model as a conceptual foundation for development of data-oriented applications. Such a data model should be based on the functional requirements gathered in the first stage and it should also be built before actually writing any application code. In the words of the author, "it's easier to change your application code than it is to change your data model" (p. 61). If well developed, a data model is crucial to the development of a good XML design that is capable of future-proofing valuable data and information. Future-proofing of data is a pragmatic concern for, in many situations, it also implies disaster-proofing.

While part 2 is concerned with different stages of application development, the third and last part (chapters 7 through 10) discusses different aspects of building application code. Note, however, that only some aspects of application building are discussed. The actual building of an application in a specific language (e.g. Perl, Java, etc.) is not within the scope of this book. Specifically, the topics covered in part 3 include related XML standards such as XSLT and XML Schema, Microsoft's SQL Server 2000, Sun's Java 2 Enterprise Edition, and further examples of integrating XML and SQL.
Generally speaking, the author has done quite well a job in explaining the significance of important concepts using the two fictitious application development scenarios. Perhaps the greatest strength of this book is the author's focus on underlying concepts. Indeed, clear conceptual understanding can, more often than not, be more important than mere knowledge of language syntax in writing efficient application code, particularly in situations where creativity is desired.

Despite the aforementioned strength, there is also some room for improvement. This reviewer believes that some of the concepts discussed in the book, such as partial decomposition and link semantics, can be more easily and better understood if they are illustrated by some computer simulations on, say, a CD. Note that in the "SQL Examples" sidebar on page 93 of the book, the author mentions a "CD that accompanies this book." However, the book that this reviewer received did not come with a CD.

In short, the author has presented some very useful and practical ideas for developing XML applications and integrating the strengths of XML and SQL. The author has also mentioned the idea of link semantics as a potential means of adding some degree of artificial intelligence to an application so that it is capable of making meaningful suggestions to end users. That is a particularly interesting idea which deserves further attention from software developers and researchers.

Overall, this book can be a useful resource for those who are somewhat familiar with XML and SQL, and who are particularly interested in integrating the two to build powerful, future-proof Web applications.

Kwan-Yau Lam is an Assistant Professor/Librarian and an adjunct Member on the Computer and Information Science faculty at Truman College, City Colleges of Chicago.

Copyright (c) 2003 by Kwan-Yau Lam. This document may be reproduced in whole or in part for noncommercial, educational, or scientific purposes, provided that the preceding copyright statement and source are clearly acknowledged. All other rights are reserved. For permission to reproduce or adapt this document or any part of it for commercial distribution, address requests to the author at klam@ccc.edu.
weakest, not because the initial studies are good at stating their case, but because save for one they are so poor as studies that one suspects Compaine has assembled a scarecrow. The first two studies are from the National Telecommunications and Information Administration. These are remarkably brief surveys and, in some cases, even the untrained can see flaws in how the data are used to justify the conclusions.

Divide then argues that despite predictions to the contrary, the gap between those social groups who have access to the Internet and those who do not is closing so that market forces are working just fine without government intervention. He argues that access to the Internet is a technology more like a subscription service (such as the telephone) than a durable good (such as radio or television). According to Divide, such services inherently take longer to adopt across all levels of society; most people who don't subscribe to such services could do so, but elect other services in their stead. The parallel to phone service is crucial, since an analysis of universal phone service is the topic of chapter five and the conclusions to that chapter are applied analogically to what follows. The potential weakness here is that one might either dismiss the analogy or -- and this seems more difficult -- refute the specific findings.

Finally, Divide argues that government interventions, such as E-rates, are poorly designed, poorly administered, and focus on immature and expensive technologies that soon become obsolete. For instance, wiring classrooms for Internet access is expensive, but at one time was the only option. Today, wireless technology is less expensive and far better suited for the classroom. Adoption of the former technology makes it more difficult to deploy the latter and, as it turns out, since the gap is presumably closing, most of the work over the Internet that could be done at school, could be done better at home.

This last argument -- that work done over the Internet at school could be done just as well at home -- assumes that one has access to the same sources at home as at school. If teachers and school libraries are relying solely on freely available material over the Internet, such a case could be made, but in fact schools typically subscribe to sources not available from home. Even so, one could still argue that the E-rates have encouraged a misplaced emphasis on immature computer and network technologies.

For librarians, Compaine's focus on access to Internet technology is largely beside the point. Librarians have, since the early days of inter-library loan, been able to distinguish between access and ownership. For librarians, whether every individual has Internet access at home is not as important as whether such access is readily available to the public. So, if a library and its branches have enough networked stations to satisfy their patrons' needs, no gap exists.

It is not, however, access to Internet technology or the freely available information on the Web that should be seen as the public good that libraries wish to serve. The part of the digital divide that has most vexed librarians is not the technology but access to information, and not just any information but commercially available information such as Lexis-Nexis or Dialog or any number of other bibliographic, full-text, or numeric databases. The information gap that librarian's face is not between those that can afford computers or access to the Internet and those that cannot, but between those who can afford access to databases that very few individuals can afford and fewer would want to save for the infrequent times they would be useful. It is this data that librarians would wish to provide as a public good just as the books, CDs, videos, and DVDs loaned out by libraries are treated as a public good. Compaine himself sites this concern (p. 107), but conflates it with mere access.

Let us assume, however, that one regards access to the Internet and its freely available content as a public good, and so libraries should invest in information technology to provide access to this to those who are underserved. In this case, Divide provides a sustained, unrelenting, and largely justified critique.
While this reviewer received the impression that Divide would have better served its goal had Compaine actually made the arguments himself, citing the contributors as sources, and placing the National Telecommunications and Information Administration studies in an appendix, the book is remarkably well organized and the articles tightly woven together to make a coherent argument. The lesson that the librarian can pull from this volume is that libraries should focus on what they have always done best: providing access to sources of knowledge to the public that would not otherwise be available. Corporations have found e-commerce to be in their interest and so have also found ready access to the Internet to be in their interest. Consequently, if libraries focus on providing simple access to the Internet, they will find themselves serving an ever-shrinking clientele. On the other hand, if libraries focus on providing the sort of information not readily found on the Internet, then they are providing a valuable public good and as a bonus have the option of integrating this service with what is readily available.

Jimm Wetherbee (jimm@wingate.edu) is Information Systems Librarian for Wingate University.

Copyright (c) 2003 by Jimm Wetherbee. This document may be reproduced in whole or in part for noncommercial, educational, or scientific purposes, provided that the preceding copyright statement and source are clearly acknowledged. All other rights are reserved. For permission to reproduce or adapt this document or any part of it for commercial distribution, address requests to the author at jimm@wingate.edu.


by Cindy Schofield-Bodt

The music I adored in high school was available to me on six- and 12-inch vinyl LPs. I lugged my record player and two milk crates full of records away to college, but when I packed up to come home four years later I moved the same music home in a shoe box to play on my cassette deck. Twenty (+!) years later I don't own a working turntable and the tape deck in my car is gone, but I still listen to 70s hits whenever I want -- off CDs purchased of the very same music. As the technology changed so did the format of the information I wanted. But what about the boxes of my grandfather's favorite dance music in the attic? The technology for playing his cylindrical recordings is long gone and I have not found one of his favorite songs, by his favorite artist, in CD format.

The myriad issues related to the preservation of digital information is the focus of Susan Lazinger's Digital Preservation and Metadata: History, Theory, and Practice. On the surface, the issues are some of the same ones faced by music lovers -- some conversion to each new format will happen but some information will remain forever locked out of reach in obsolete formats as technology continues to change. According to Lazinger, "intellectual preservation of electronic data, as opposed to preserving the medium on which information is stored, is an issue that originated with the creation of electronic data" (p.6). In the environment of the printed text, preserving the artifact preserved the intellectual content. The "fixity" of text is taken for granted in the printed world. Electronic texts can be so easily manipulated, revised, improved and destroyed that there is no such assurance of permanence. The electronic information age introduced new preservation requirements and challenges. The proliferation of information, the uncontrolled accumulation of data, the realization that catastrophic electronic data loss has already been faced by the
U.S. Census Bureau and the U.S. military, the prevalence of electronic data designed to prevent copying and a myriad of other issues have experts racing to organize, standardize, institutionalize and empower caretakers of our digital heritage.

Lazinger writes that her book was originally intended as a collation of the latest theories and projects in the area of electronic publishing, but in the three years it took to write it the book became instead a "history of digital preservation from its inception to the beginning of the 21st century" (p. xx).

Digital preservation has become an issue because of the fluidity of the medium and the precarious nature of the intellectual content itself. Data can be easily accidentally or inadvertently destroyed or tampered with, hardware and software quickly becomes obsolete and there is no empowering mechanism that would allow institutions to become "official" caretakers of our electronic resources. The printed word may have continued to exist in spite of benign neglect of our ancestors but the same grace will not be afforded our electronic endeavors. Today a 400-year-old book can be read more easily than a 15-year-old data disk. "If we are to pass on our digital heritage to our descendents, as our ancestors passed on their printed heritage to us, preservation decisions will have to be integrated with the creation process itself" (p. 15).

Chapters 2 and 3 crystallize two of the main talking points in any discussion of electronic preservation: "What electronic data should be preserved?" and "Who should be responsible for electronic preservation?" The answer to "what" is not so different as the answer when regarding traditional information formats, but "new dimensions that digital data adds to preservation-fluidity and a dynamic nature-in fact raise additional questions such as which version of a resource is the 'genuine' one" (p.17). When multiple versions and updates are considered the question of "what" to preserve also gets extended into a question of "how much" of something to preserve. In addition, there are differing considerations for information that is "born digital" vs. something that has been digitized but originated as a print source. The answer to what should be preserved is not one that Lazinger gives a pat answer to but she directs the reader to collection policies and selection criteria of individual research libraries as well as the "List of Quality Selection Criteria: A Reference Tool for Internet Subject Gateways," written and distributed by the English-based Telematics for Research project DESIRE" (p.39).

Although stakeholders always include the creators of digital objects -- individuals, institutions, or organizations -- other stakeholders in the preservation mix include publishers, distributors, systems administrators, librarians, archivists and end users. While all of the stakeholders are united in their interest in adding to or making use of the value of digital information, the roles of the stakeholders have not yet been clearly defined. So far, libraries have been "in the forefront of research into and implementation of the beginnings of far-reaching preservation policies for digital resources" (p.64). There is no comprehensive system in place to collect U.S. government files and it is not clear that the "legal deposit" provision that obliges publishers to deposit copies of their publications in libraries in the countries in which they are published, applies to non-print publications. Obviously, there are huge gaps in what digital material is intentionally collected and initiatives are just beginning to be developed for harvesting online publications in the U.S. and Europe. In the United States, OCLC and RLG are deeply involved stakeholders and are making great strides in the organization of digital preservation projects.

Chapter 6 provides an excellent overview of "Models for Syndactic and Semantic Interoperability," and is so titled. Metadata -- referring to data that describes networked electronic resources -- "increases the odds that a user will be able to retrieve appropriate information and assess its usefulness and availability" (p. 174). Metadata is integral to the preservation process, as it is a complete record that includes with the description of the document the information necessary for managing and preserving the information being described.
Some metadata is automatically generated when a networked document is created; a second type is created manually. It is stored as a part of the resource itself according to standards that make it useful to various resource communities which are described in detail in the chapter.

The data's equal twin in preservation considerations is the physical mechanism upon which the data is recorded. "Technologies have been developed that are designed to handle digital media of all types" (p. 190), and the underlying principles are discussed in chapter 7. The OpenDoc Standard Interchange Format (Bento specification) involves "wrapping" the data in "object containers," thus storing the "digital object and the metadata together to explain both the content and how to use it at some future time." Lazinger goes on to describe Open Media Framework (OMF) Interchange and the Warwick Framework and the many developments in the Universal Preservation Format (UPF) aimed at ensuring accessibility of many data types into the future. Unifying standards are still being developed and formalized as the metadata community evolves with the introduction of each new framework.

Chapter 8, "From Theory to Reality: Selected Electronic Data Archives in the United States" gives an overview of many of the current digital preservation projects being undertaken in the United States and around the world. The projects vary in scope, intended audience, content and format. The first half of the chapter focuses on cultural heritage digitization projects; the second half examines archives that house social science data. These are not intended as complete lists but rather as models of digitization technology. The various cooperatives and consortia highlighted are often working together to create workable standards and to secure funding for local, regional, and national preservation initiatives. The Chapter entitled "Further Reality: International Digital Cultural Heritage Centers and Sites and Electronic Data Archives" goes on to cite some of the international digital projects that are beginning to showcase the wealth of cultural treasures around the world.

Lazinger includes an extensive bibliography, including useful Web sites that are examples of preservation, as well as resources about digital preservation. A detailed index provides access to the material presented. This book is highly recommended for anyone in the information field. It brings together the conundrums as well as some of the resolutions to the digital preservation quandary. The historical references to preservation development will continue to shed understanding on the issues for some time; the current solutions to the dilemmas faced by today's program designers are an important record and will serve as models for future development even after today's relevancy is gone.

Cindy Schofield-Bodt (Schofieldbc1@southernct.edu), Librarian, Southern Connecticut State University, is currently the Buley Library Technical Services Division Head. She has taught cataloging and acquisitions courses as an adjunct at the SCSU Graduate School of Library And Information Science.

Copyright (c) 2003 by Cindy Schofield-Bodt. This document may be reproduced in whole or in part for noncommercial, educational, or scientific purposes, provided that the preceding copyright statement and source are clearly acknowledged. All other rights are reserved. For permission to reproduce or adapt this document or any part of it for commercial distribution, address requests to the author at Schofieldbc1@southernct.edu (mailto:Schofieldbc1@southernct.edu).

As more and more educational institutions are providing distance education courses and online learning, even if one does not teach an online or distance education course, there is a need to understand this environment to support students in these programs. Carolyn Nash, the administrator of the Central Massachusetts Regional Library System in Shrewsbury Massachusetts and manager of its Web-based distance education program, and Linda Braun, an educational consultant and teacher in the University of Maine Library and Information Technology Distance Education Program, attempt to do just that in their book, The Browsable Classroom. Even though the title implies that this is written for librarians, the book would be of value to anyone who wants to gain a very basic understanding of the distance education environment, understand how it differs from traditional instructional environments, obtain practical suggestions, and find examples of existing distance education courses and programs.

The book is divided into seven brief chapters on different topics. Each chapter stands alone, exploring each topic, thus allowing a reader to pick areas of interest to read and discover further using the references to the cited resources at the end of each chapter. In the first chapter, one will find discussions of types of online environments, synchronous and asynchronous, and their potential components. The second gives us a look at successful programs so we can learn from others' experiences. Though a majority of the programs discussed are library programs such as TILT and FALCON, these can be used as models for other learning environments. The third chapter discusses support issues libraries face and the fourth takes us through the design process and considerations when developing a course or program. Chapter five addresses adapting traditional pedagogy to an online environment as traditional face-to-face courses cannot be put on the Web without careful evaluation and modification of delivery. Chapter six deals with the problems students face in this environment and the last chapter concludes with a look at the future. The glossary and index also make this book a useful reference.

The book was enjoyable to read as the authors pull the reader into the world of online learning, making it personal and inviting. The tone is set in the preface, where the reader is asked to: "Picture this: It's late at night and you're sitting at your computer in your pajamas sipping a cup of tea. You log on to the Web site posted for the course you enrolled in by e-mail." Many of the chapters make you feel as though you are there and involved in the actual process or situation being discussed. Including quotes from interviews and sharing the experiences others have had working in this new environment, enhances this engaging and easy-to-read writing style.

As an introduction, this book accomplishes the task very well as the authors offer a balanced presentation of basic theory, short concise explanations, practical examples, and suggestions devoid of technical jargon. One will find useful examples and Web addresses to existing online courses as well as some sample forms, scripts, charts, and checklists for different phases and components of the design process and course presentation. Although other modes of distance education are mentioned, the book predominantly presents distance educational environments that use the Internet as their mode of delivery. As I have taken coursework in distance education and instructional design, I would advise that others should not assume, after reading this book, they will know everything they need to know to create a successful course or program. This is, as the title states, an introduction and is useful for anyone trying to gain a basic understanding of this new environment. If a reader is going to develop an online course or program, he or she should read further about online pedagogy and instructional design concepts and principles found in other books, such as:
Two classics on instructional design:


by Maribeth Manoff

In the reading I have done about XML (Extensible Markup Language), I have amalgamated the different definitions of XML into one of my own: It starts with a simple definition of HTML (Hypertext Markup Language). HTML is a markup language, specifying tags and attributes that define how text will be displayed in a Web browser. HTML is a subset of the powerful and complex SGML (Standard Generalized Markup Language). XML is also a subset of SGML, but rather than being a markup language, XML provides a set of tools for programmers to create their own markup languages. With XML, users can create tags and attributes that describe the content of a document, and then style sheets can be used to define how that content is to be displayed. This functionality allows XML to overcome one of the major problems inherent in HTML, the inability to separate content from display. The World Wide Web Consortium's (W3C) "XML in 10 points" describes the power of XML: "XML makes it easy for a computer to generate data, read data, and ensure that the data structure is unambiguous." Although I was satisfied with this definition and with the potential of XML to transform some of the work that I do, I was looking for more.
First, there were other acronyms and specifications I had seen mentioned along with XML: DTDs (Document Type Definitions); Namespaces; XLink; XPointer; Schemas. There was also the fact that I learn best by doing.

Was XML a book that could help me put some of these pieces together? Did it have the hands-on exercises and clear exposition that could have me actually using XML by the time I was done with it? The fact that this book is part of an Addison-Wesley series called The Nitty Gritty Programming Series gave me hope that these expectations could be met.

Part I of the book is entitled "Start Up!" It begins with a cursory explanation of the relationship between SGML, HTML and XML, and then moves into a chapter containing descriptions, with screen shots, of some browsers and their rendering of XML, along with similar treatment of several XML editors.

Next comes the real nitty gritty: Chapters 3 through 8 are Descriptions and exercises leading to a comprehensive view of some XML basics. Chapter 3 is entitled "The essentials of XML." It starts with an exercise common to many programming texts, sending the words "Hello World!" to the screen. A simple XML document and Cascading Style Sheet (CSS) are used to complete this task. The reader can follow along, using a text editor and XML-compliant browser. A more complex example is then introduced, that of a personnel file. Step-by-step instructions are provided for taking data that could make up a typical personnel file and converting that data into XML.

Chapter 4 describes in detail how to write a Document Type Definition (DTD), which sets down "rules for the elements and attributes within an XML document" (p. 47), again using the personnel file example. Chapter 5 goes into more detail about Cascading Style Sheets, and how they are used in conjunction with XML, then goes on to another method of formatting XML data, Extensible Stylesheet Language Transformation (XSLT). Chapter 6 describes the linking capabilities available in XML, a standard known as XLink. As browsers do not yet support XLink, this chapter is brief, not having any way to extend the explanation to real examples. Part I is completed by another brief chapter describing a way to write HTML as XML code, known as XHTML (Extensible Hypertext Markup Language).

Part II of XML is called "Take that!" It contains a set of "quick reference guides" for XML, XSLT, and CSS. Another reference guide to HTML, completes Part II.

Part III, "Go ahead!," introduces the XML application WML (Wireless Markup Language) which is used to create Internet pages for mobile phones. Again the "Hello World" exercise is used, followed by the creation of several other pages that can be viewed in a WAP (Wireless Application Protocol) browser. One more quick reference guide for WML completes this part. Finally, there are two appendices: a glossary, and a set of useful links to help the reader "keep on the ball" with development of XML.

In many ways, XML met the hopes that I had for it. In particular, the chapters covering the basics of XML, DTDs, and style sheets were very good. The way these chapters used and expanded on the personnel file exercise met two important goals. First, they covered a good number of XML basics such as elements, attributes, valid XML, and namespaces using a hands-on approach, so that I could put the concepts together as I worked through the exercise. Second, the power of XML became evident. XML tags such as "<Department>" and "<Person>" describe the content of the personnel file in an intelligible way. This content can be displayed in many different ways using various style sheets. The coverage of XLink was also well done, in terms of explaining basic concepts in an intelligible way, but the chapter on XHTML was not quite as helpful. The W3C describes XHTML as the "successor to HTML," yet there was no explanation of why or how that might come about. The authors briefly outline the differences between XHTML and HTML, but are not clear about the advantages to using one over the other.
The reference guides for XML, CSS, and XSLT, and the chapter on WML, I'm sure will be useful as I go on to discover and try more XML applications. The only part of the book that confounded me was the extensive reference guide to HTML. There were some useful lists here of HTML versions and which browsers work with which tags and attributes. In the scheme of things, though, it was disappointing to have so much space devoted to this. I would have rather seen some help with XML concepts that are not covered, such as XPointer and Schemas.

All in all, I would recommend this book to anyone who has a basic understanding of what XML is, and who would like to get started trying to use this exciting new language.

Reference:


Maribeth Manoff, Systems Librarian, University of Tennessee, Knoxville.

Copyright (c) 2003 by Maribeth Manoff. This document may be reproduced in whole or in part for noncommercial, educational, or scientific purposes, provided that the preceding copyright statement and source are clearly acknowledged. All other rights are reserved. For permission to reproduce or adapt this document or any part of it for commercial distribution, address requests to the author at mmanoff@utk.edu.

About TER