EDITED BY WILLIAM E. KASDORF

The Columbia Guide to **Digital** Publishing

Columbia University Press • New York

Columbia University Press
Publishers Since 1893
New York Chichester, West Sussex

Library of Congress Cataloging-in-Publication Data
The Columbia guide to digital publishing / edited by William E. Kasdorf.
p. cm.
Includes bibliographical references and index.
ISBN 0-231-12498-8 (alk. paper) — ISBN 0-231-12499-6 (pbk. : alk. paper)
1. Electronic publishing I. Kasdorf, William E.
Z286 .E43C65 2003
070.5'797—dc21 2002041462

Columbia University Press books are printed on
permanent and durable acid-free paper.
Printed in the United States of America
c 10 9 8 7 6 5 4 3 2 1
p 10 9 8 7 6 5 4 3 2 1

Contents

Credits *xiii*
Contributors *xiv*
Preface *liii*

Chapter 1. Introduction: Publishing in Today's Digital Era *1*

William E. Kasdorf, President, Impressions Book and Journal Services, Inc.

Today, virtually all publishing is digital to some extent, whether content is delivered electronically or in print. But every publishing family is digital in its own way. Some, like journals and reference publishers, have moved so far toward electronic publishing that they are beginning to abandon print. Others, like magazines and catalogs, have focused more on digital production technology and, with newspapers, on integrating workflow for print and online publication. E-books still seem tentative, although they present big advantages, especially to publishers of textbooks and scholarly books. This introduction examines these different sides of publishing today to provide some perspectives on the many aspects of publishing in the digital era that this Guide discusses in depth in subsequent chapters. This overview ends with some suggestions of important technologies that deserve special attention from publishers today.

1.01 Digital publishing is both a given and a goal *1*
- (i) The real revolutions have already happened *2*

1.02 Various publishers, various solutions *3*
- (i) Journals: the proving ground for electronic publishing *3*
  - – The problem with print *4*
  - – Awkward economics *4*
  - – Moving from print to electronic publication *5*
  - – Pioneering SGML *5*
  - – Pricing and delivery models *6*
  - – Making access ubiquitous *6*

- DOI and CrossRef *7*
- It's real *7*

(ii) Magazines and catalogs: where graphics became digital *7*
- PostScript + PC = DTP *8*
- A whole new way to work *8*
- It's all about print *8*
- What about electronic publishing? *10*
  - Print production tools for online uses *10*
- Content management and digital asset management *11*

(iii) Textbooks: a tough transition to electronic publishing *11*
- The obstacles to electronic publishing *12*
- Custom publishing and coursepacks *13*

(iv) Scholarly monographs: saved by print-on-demand and e-books *15*
- Print-on-demand *15*
- E-books *16*

(v) Trade books: the rights battleground *18*
- E-book trials, failures, and partial successes *18*
- Wrangling over rights *18*
  - Authors vs. publishers *19*
  - Publishers vs. readers *19*
  - The perfect copy problem *20*
  - Learning how to *manage* digital rights *20*

(vi) Reference: where metadata matters most *21*
- Isn't it all free on the Web? *21*
- Metadata makes it really work *22*
- Adding intelligence adds value *23*

(vii) Newspapers: putting it all together *23*
- Digital technology makes it easier, not more complicated *24*
- Beyond managing content to managing workflow *24*

1.03 What next? *25*

(i) Some givens *25*
- The Internet *25*
- The Web *25*
- XML *26*
- PDF *26*
- DOI *26*
- Unicode *26*

(ii) Some works in progress *27*
- The personal computer *27*
- E-books *27*

– DRM 27
– Linking 27
– Metadata 28
– JDF 28
– Digital printing 28
– Broadband 28
– Accessibility 28
(iii) Some future prospects 29
– The Semantic Web 29
– Web services 29
– XSL-FO 29
– E-paper 29
– XMP 30
– SVG 30
– Taxonomies 30
– Mapping information 30
1.04 Dive in! 30

Chapter 2. The Technical Infrastructure 32
Chris Biemesderfer, Seagoat Consulting

This chapter discusses the basic hardware and software that provide the infrastructure for digital publishing. Becoming familiar with these components will help those involved in digital publishing to better understand technical issues and use tools and techniques more effectively.

2.01 Overview 32
2.02 The basics of computer architecture 32
(i) The anthropomorphized computer 32
– Thinking: processors, memory, and the processing environment 33
– Showing: display technology 33
– Saving: storage technology 34
– Talking: communication and data transfer 34
(ii) A few basic terms and definitions 34
– Bits and bytes 34
– Input and output 35
– Digital vs. analog 35
2.03 The processing environment 36
(i) Operating systems 36
– Unix 36
– Apple Macintosh 37

– Microsoft Windows *38*
– Other operating systems *39*
• Large: mainframes and supercomputers *39*
• Tiny: handheld and embedded *40*
(ii) Application software *41*
– Concurrent calculation: spreadsheets *41*
– Storage, search, and retrieval: databases *42*
– Manipulation of text: word processing and formatting *43*
– Manipulation of nontext: graphics, design, and image/audio processing *43*
– Combinations: accounting packages, content/asset management, and more *44*
• Accounting *44*
• CAD: computer aided design *44*
• Content management *45*
2.04 Display *45*
(i) Terms and definitions *45*
– Bitmaps and rasters *45*
– Pixels and resolution *46*
– Monochrome, grayscale, and color *46*
– Antialiasing *46*
(ii) Overview of display technology *47*
– CRTs: cathode-ray tubes *47*
– LCDs: liquid crystal displays *47*
– Touch sensitive displays *48*
– Electronic ink *48*
– Video signals *48*
2.05 Data storage *49*
(i) Storage devices *49*
– Magnetic media *49*
– Optical media *49*
– Hybrid media *50*
– Media and format *50*
(ii) Disk storage *50*
– Fixed disks *51*
– Removable disks *51*
– Optical disks *51*
• Compact discs and CD-ROM *52*
• Recordable and rewritable CD: CD-R and CD-RW *52*
• DVD *52*
(iii) Tape storage *52*

2.06 Data communications 53
(i) Local device communication 53
– Communications channels 53
• Serial channels 53
• Parallel channels 54
– Common examples of device communications standards 54
• RS-232 54
• FireWire 54
• USB 54
• Parallel printers 54
• SCSI 55
• GPIB 55
(ii) Network communications 55
– The client/server protocol 55
– Network topology 56
(iii) Linking computers: LANs and WANs 56
– LAN: local area networks 56
• Wireless LANs 57
– WAN: wide area networks 57
• Dial-up analog circuits and modems 57
• Digital circuits 58
• Circuit speed designations 59
(iv) Communication rules: protocols 60
– Networking protocols 61
– Application protocols in TCP/IP 62
• Terminal and file sharing services 62
• Information exchange 62
• Application services and discovery 63
2.07 Additional interesting resources 63

Chapter 3. Markup: XML & Related Technologies 65
William E. Kasdorf, President, Impressions Book and Journal Services, Inc.

Markup enables the various parts and features of a given set of content to be distinguished and named. It provides a way to label, describe, and delimit these in a publication so that processing systems can tell them apart and know how they relate to each other. Markup languages are used to define specific markup schemes. In the past, markup languages were typically proprietary and used only by specialists. The Web gave rise to one of the simplest and most widely used markup languages ever devised, HTML, and also to one of the most flexible and powerful: XML, the Extensible Markup

Language. After a brief overview of earlier markup languages, this chapter focuses on the technologies in the XML family— XML itself, and related standards for defining, styling, linking, transforming, and annotating—that provide the foundation for digital publishing today.

3.01 Overview *65*

- (i) What is markup? *65*
  - – Visual cues *66*
  - – Proprietary markup *66*
  - – Rigorous vs. ambiguous markup *68*
    - • Formatting ambiguity *68*
    - • Structural ambiguity *68*
    - • Missing codes *69*
  - – Separating structure and appearance *70*
    - • Markup for structure *71*
    - • Markup for meaning *72*
- (ii) A brief history of structured markup standards *72*
  - – GenCode and GML *73*
  - – SGML: the Standard Generalized Markup Language *74*
  - – HTML: HyperText Markup Language *75*
  - – XML: the Extensible Markup Language *75*
- (iii) Open standards *76*
  - – SGML and ISO, the International Organization for Standardization *76*
  - – XML and the W3C *77*
  - – OeB and the Open eBook Forum *78*

3.02 HTML: HyperText Markup Language *79*

- (i) Parts of an HTML document *80*
  - – The header *81*
  - – The body *81*
  - – HTML character entities *82*
  - – HTML editing tools *82*
- (ii) Web programming *82*
- (iii) Accessibility *83*
- (iv) Formatting and style *84*
  - – CSS (Cascading Style Sheets) *85*
  - – XHTML *85*

3.03 XML: the Extensible Markup Language *85*

- (i) Overview *85*
  - – XML is a metalanguage *87*
  - – XML is also a family of standards *88*

- XML is used for other standards *89*
- XML and SGML *89*
- XML and HTML *91*

(ii) Using XML for structured markup *92*
- Document analysis *92*
  - Scope *92*
  - Granularity *94*
  - Functionality *95*
  - The politics of document analysis *96*
- Document Type Definitions (DTDs) *97*
  - Elements *97*
  - Elements delineate structure *98*
  - Attributes *100*
  - Empty elements *101*
  - Entities *101*
  - Defining rules and parsing *104*
  - An example of a simple DTD *105*
- The role of standard DTDs *108*
  - ISO 12083: it's all about structure *109*
  - TEI: a wealth of options *111*
  - MEP: stages of enrichment *113*
- Schemas *114*
  - XML schema languages *115*
  - Datatyping *116*
  - Two schema examples *116*
  - RELAX NG *116*
  - XML Schema *118*
- The document instance *120*
  - Well-formed vs. valid XML *121*
- Namespaces *122*

(iii) Defining the appearance of XML documents *123*
- Converting to HTML for Web display *124*
- Cascading Style Sheets (CSS) *124*
- XSL, the Extensible Stylesheet Language *126*
  - XSLT and XSL-FO *126*
  - XSL-FO is a specification language, not a formatting engine *128*
- Fonts and character encoding *128*
  - Unicode *130*

(iv) Transforming XML documents *131*
- Manual conversion with text editors *132*

– Scripting *133*
– Validating editors *134*
– XSLT: XSL Transformations *135*
(v) Linking in XML *138*
– XLink: the XML Linking Language *140*
– XPath: the XML Path Language *142*
– XPointer: the XML Pointing Language *144*
(vi) Metadata *145*
– RDF: the Resource Description Framework *146*
– XMP: the Extensible Metadata Platform *148*
– Metadata vocabularies in publishing *149*
• Dublin Core *149*
• ONIX *150*
• CrossRef *151*
3.04 Communication, cooperation, collaboration *152*

Chapter 4. Organizing, Editing, & Linking Content *155*
John Strange, Group Production Director, Blackwell Publishing

The production and distribution of content are two of the most important functions carried out by publishers. Historically, these publishing functions had only to concern themselves with how to produce and distribute physical items: books, magazines, and journals printed on paper. However, since the mid-1990s publishers have recognized that in order to meet the demands of digital publishing, their publishing processes require radical rethinking. The traditional print requirements must continue to be fulfilled, but the additional demands of digital publishing must also be met. The content must be available electronically—usually on-line. It must be consistently produced to a high standard, available at the same time or ahead of print, and it must provide maximum functionality—which often means linking both internally and externally, and sometimes means including nonprint content. Of course, all of this must be achieved at minimum added cost to the publisher.

4.01 Overview: the transition from traditional to digital publishing *155*
4.02 Structuring content *156*
(i) XML in the digital production process *156*
– XML-out *156*
– XML-in *157*
– Benefits of XML-in vs. XML-out *158*
(ii) Choosing a DTD *158*
– Bespoke (or Base) DTDs *159*

• Sfx: an example of an open linking solution *176*
• DOI/CrossRef and OpenURL/SFX *176*
4.06 Conclusion *176*

Chapter 5. Data Capture & Conversion *179*

Mark Gross, President, Data Conversion Laboratory

This chapter discusses the ins and outs of converting and upgrading textual data and related information from all kinds of formats into markup languages for publishing purposes. While the issues and approaches apply to all markup languages, particular emphasis is given to the issues of converting documents into XML, SGML, and HTML. The chapter looks at the issues surrounding the capture of data from sources such as paper and microfilm, word processors, publishing systems, PDF, and ASCII—the main issue being the need to untangle content from structure.

5.01 Overview: Entering a world of structure *179*
(i) A step-by-step approach *180*
– Define aims *180*
– Determining the appropriate level of tagging *180*
– Analyze your data *181*
– Data capture *182*
• Extraction of text *182*
• Determining text flow and structure *182*
• Obtaining nontext objects *183*
• Establishing relationships between text and objects *183*
– Creating your target format *183*
– Output *183*
5.02 Untangling content from structure *184*
(i) Why is it important? *184*
(ii) Implicit vs. explicit information *185*
– Identifying ambiguity *185*
5.03 Where does your data come from? *187*
(i) Legacy content *187*
– Nondigital sources *187*
• Paper and microfilm *187*
• Key-entry processes *188*
• OCR *189*
• How much accuracy do you need? *190*
– Digital sources *190*
• Word processors *190*

– Standard DTDs *159*
– Granularity *160*
4.03 The impact of digital publishing on traditional publishing models *160*
(i) Case study: journal publishing in the digital world *161*
– Supplementary content *161*
• Supplementary content: recommendations *161*
– "Disaggregation" of content *162*
• "Disaggregation" of content: recommendations *163*
4.04 Information about content: metadata *164*
(i) What is metadata and why is it important? *164*
(ii) Defining metadata *164*
– Metadata vocabularies *164*
– "Core" metadata *165*
– Examples of metadata vocabularies *165*
• DCMI *165*
• ONIX *165*
• MARC *166*
(iii) Structuring metadata *166*
– Resource Description Framework (RDF) *166*
– Exchanging metadata *166*
(iv) The Semantic Web *167*
4.05 Linking *168*
(i) Intradocument links *168*
– Intradocument links: recommendations *169*
(ii) Interdocument links *170*
– Interdocument links: recommendations *170*
(iii) Extradocument links *171*
– URL (Uniform Resource Locator) *171*
– FTP (File Transfer Protocol) *171*
– DOI (Digital Object Identifier) *171*
– E-mail (Electronic Mail) *172*
(iv) Extradocument links: recommendations *172*
(v) Reference linking *172*
– CrossRef *172*
• How CrossRef linking works *173*
• CrossRef linking requirements *173*
• How the CrossRef process works *174*
– Abstracting services *175*
– OpenURL *175*

– Packaging and delivery *209*
– Exception handling mechanisms *209*
(v) Where do you go from here? *209*
– The six keys to successful conversion *209*
5.05 Analysis issues *210*
(i) Text and characters *210*
– Fonts, special character encoding, Unicode *210*
– Accents and diacritics *211*
– Non-Latin alphabets *211*
(ii) Tagging of paragraphs and heads *212*
(iii) Cross-references and linking *212*
(iv) Lists *212*
(v) Tables *213*
(vi) Graphics *214*
– Raster *214*
– Vector *215*
(vii) Indexes *215*
(viii) Bibliographic information *216*
(ix) Footnotes *217*
(x) Math *217*
5.06 Summing up *217*

Chapter 6. Composition, Design, & Graphics *219*
Thad McIlroy, President, Arcadia House

Publishing in the digital era, to an overwhelming degree, still means publishing in print. More often than not what is published electronically today first started as print. While much of the energy in digital publishing has shifted to the electronic conveyance and display of information, the bulk of the economic activity still centers on print-based publishing. Electronic publishing is the future of the industry, but print publishing pays today's bills. Today's digital tools, technologies, and techniques are as often used in print as in electronic publishing. Mastering print publishing demands knowledge of a broad range of processes and technologies surrounding composition, design, and graphics. Print publishing has reached the intersection of 500 years of analog craft and a new era in digital computer graphics. Publishing practitioners require an appreciation of both. This chapter is a broad overview of the processes and technologies that comprise the practice of print publishing in the era of digital publishing. We focus on the digital technologies that are moving the process forward into the future, *and the traditional craft-based skills still important in the digital era that optimize quality in everyday work.*

6.01 Overview *219*

6.02 Text, graphics, and page layout: The three elements of a page *220*

6.03 Design vs. production *221*

6.04 Three key technologies *221*

6.05 PostScript: the language of print publishing *222*

- (i) Competition *223*
- (ii) Technical aspects *223*
- (iii) PostScript versions *224*
- (iv) Type in PostScript *224*

6.06 PDF—Adobe's Portable Document Format *224*

6.07 Typography *227*

- (i) PostScript *228*
  - – Type 1 fonts *229*
  - – TrueType *229*
  - – Multiple Master fonts *230*
  - – OpenType *230*
- (ii) Managing fonts *231*
  - – Screen fonts *231*
  - – Printer fonts *231*
  - – Adobe Type Manager and other font management software *231*
- (iii) Font troubleshooting *232*
  - – Characters are okay on the screen, but don't print correctly *232*
  - – Characters are italicized or bold on screen, but not when printed *232*
  - – Characters are bitmapped on screen (i.e., they are not drawn smoothly, but with big squares) *233*
  - – When you select a font, another one appears—or it appears, but with odd spacings or leadings *233*
  - – Text is printed in Courier instead of the PostScript font on your screen *233*
- (iv) Font embedding *234*
  - – Identifying the foundry *234*
    - • Use ATM Deluxe's sample window *234*
    - • View the AFM file *234*
    - • Use the Get Info command (Macintosh only) *234*
    - • Use the TrueType sample sheet (Windows only) *234*

- Use the Microsoft Font Properties extension (Windows only) 235

– Contacting a foundry 235

(v) Controlling type 235

– Typefaces 235

– Measuring type 236

– Spacing 236

- Optical spacing 236
- Letterspacing 237
- Kerning 237
- Tracking 237
- Leading 238
- Line length 238

– Type (point) size 239

– Alignment 239

– Fixed spacing 240

– Hyphenation and justification 240

– Word spacing 241

(vi) Special typographic situations 241

– Symbols 241

– Equations and math 242

– Accents 243

– Quotes 243

– Dashes 244

- Hyphen 244
- En dash 244
- Em dash 244

– Ligatures 244

– Rules 245

– Ellipsis 245

– Small caps 246

6.08 Graphic types and file formats 246

(i) Graphic types vs. file formats 247

(ii) Graphic types 247

– Bitmaps 248

– Vector graphics 248

– Fonts 249

– Pages 249

(iii) File formats 249

– Types of file formats 250

(iv) File compression *251*
- Compression software *252*
(v) Cross-platform file issues *252*
- File format specifics *253*
  - TIFF *253*
  - TIFF/IT *253*
  - JPEG *254*
  - PICT *254*
  - EPS *255*
  - PostScript *255*
  - PDF *256*
- Special graphics formats *256*
  - OPI *256*
  - DCS *257*
(vi) Imaging and halftones *257*
- Continuous tone *257*
- Dots, spots, samples, and pixels *258*
- Halftone variables *259*
  - Screen ruling *259*
  - Screen angles *260*
  - Dot shapes *261*
- Dot gain *261*
  - Dot gain compensation *262*
- AM vs. FM screening *263*
- Duotones *264*

6.09 Color *264*
(i) Thinking about color *264*
- Additive and subtractive color *265*
(ii) Color systems and color gamuts *265*
- Color in scanners and monitors: RGB *266*
- Printing color: CMYK and spot color *266*
  - Process color *266*
  - Defining process colors *267*
  - Spot colors *267*
  - Using spot colors *268*
  - Varnish *268*
- High-fidelity color *268*
- Four-color transitions *269*
(iii) Color reproduction in print *270*
- Gray balance *270*
- Tone reproduction *270*

– Tone compression 270
– Color correction 271
– Selective color correction 271
– Image sharpness 271
– Memory colors 271
(iv) Proofing 272
– Preproofs and contract proofs 273
– Remote proofing 273
– Printing conditions 274
– Disadvantages of digital proofing 274
– Acceptance of digital contract proofs 274
– Viewing conditions 275
• The viewing booth 275
– The human factor 275
(v) Color management systems 275
– What are color management systems? 276
• Characterization 276
• Calibration 276
• Transformation 277
– Device-independent color 277
– Implementing color management 277
– Monitors and color management 278
– Color management beyond print 278
6.10 Page production 278
(i) Desktop publishing 279
– QuarkXPress 279
– Adobe InDesign 280
– Supplementing desktop systems 280
(ii) Other page production systems 280
– High-end composition systems 281
– Tech doc systems 281
– Ventura 282
– TeX 283
6.11 Image capture and image processing 284
(i) Scanning 284
– High-end scanners 284
– Transparencies vs. reflective art 285
(ii) Digital cameras 285
– Resolution 286
– Color depth 286
– ISO values 286

(iii) CCDs *287*
(iv) Choosing a scan resolution *287*
- Continuous tone images *287*
- Scanning line art *288*
- Scanning for the screen *288*
- RGB vs. CMYK *289*
- Image reuse and archiving *289*
- Scanner tips *289*
(v) Adobe Photoshop *290*
(vi) Clip art and digital photo libraries *290*
6.12 Workflow *290*
(i) Component files *291*
(ii) Consolidated files *291*
(iii) Editability *291*
(iv) Obstacles to workflow automation *292*
(v) Workflow automation *292*
(vi) Workflow design *292*
(vii) Networking *293*
- Servers *294*
- Open Prepress Interface (OPI) servers *294*
- DCS and OPI *294*
(viii) Trapping *295*
(ix) Trim, imposition, and binding *296*
(x) Preflight *296*
- Common prepress file problems *297*
(xi) PDF and prepress workflows *298*
- Acrobat and PDF versions *299*
- PDF workflows *299*
  • PDF is (somewhat) bulletproof *299*
  • PDF is compact *299*
  • PDF is efficient *299*
- Establishing a good PDF workflow *300*
  • Problem one: garbage in, garbage out *300*
  • Problem two: fonts still a headache *300*
  • Problem three: distiller's ugly interface *301*
- High-end systems *301*
- PDF/X *302*
  • PDF vs. PDF/X *302*
  • Which version? *303*
  • TIFF/IT vs. PDF/X *303*
- For more on PDF *303*

(xii) Digital Asset Management (DAM) *304*
- What is digital archiving? *304*
- The dual value of digital data *305*
- Media formats *306*

(xiii) The future of print workflows *306*
- The Digital Smart Factory and CIM *306*
- JDF *308*
- PrintTalk *309*
- PPML *309*
- Other initiatives *310*
- Process control *310*
- On-press color control *311*

(xiv) Print buying on the Web *312*

6.13 Printing processes *314*
(i) Offset printing *314*
- Types of offset presses *315*
- Inking system *315*
- Waterless offset *315*

(ii) Letterpress *316*
(iii) Flexography *316*
(iv) Gravure *317*
- Uses of gravure *319*
- Disadvantages of gravure *319*

(v) Screen printing *320*
(vi) Digital printing *320*
- Large format inkjet *321*

6.14 Working with printers *322*
6.15 Resources *322*
(i) File formats *322*
(ii) Imaging *323*
(iii) Understanding color management *323*

Chapter 7. Accessibility *325*
Frederick Bowes, III, Electronic Publishing Associates

Digital publishing technologies offer publishers and other content providers a new dimension in shaping their publications. Well-constructed digital content can be configured to be used with a wide range of accessibility tools, including specialized software such as text-to-speech programs, designed to help people with disabilities gain access to published materials that would otherwise be unavailable to them. This has not been lost on disability

advocates, who have driven recent federal and state legislation that puts increasing pressure on publishers to provide their products in formats accessible to people with a variety of disabilities. This chapter will inform the reader of key issues, problems, and opportunities in content accessibility and what publishers must do, operationally, to meet accessibility needs. It also contains strategic suggestions about factoring accessibility requirements into product and business plans.

7.01 Overview *325*

- (i) Inaccessibility is costly in the Information Age *326*
  - – Education and knowledge are essential tools in the 21st century *326*
    - • The need for knowledge workers *326*
    - • High-stakes tests in schools define minimum competence *327*
    - • Lifetime learning requires accessibility at every age *327*
  - – Economic costs resulting from content accessibility barriers *327*
    - • Lost contribution of potentially productive workers *327*
    - • Higher costs of unemployment *328*
    - • Lower productivity from under-skilled workers *328*
    - • Accommodations are expensive *328*
  - – Social cost of barriers *328*
- (ii) The climate is changing *329*
  - – Diversity is becoming law *329*
  - – Strong disability laws and regulations taking hold *330*
  - – Chafee amendment to the Copyright Act *331*
  - – High-stakes student testing to validate competence is proliferating *331*
- (iii) Configurable content conquers most accessibility issues efficiently *332*
- (iv) Digital technologies now make configurable content a practical paradigm *333*
  - – XML affords a flexible, cost-effective solution *334*
- (v) Digital does not mean accessible *334*
  - – Accessibility requires careful tagging *335*
  - – Tagging alone is not enough *335*
    - • Depth of detail *336*
    - • Tables, charts, and graphs *336*
    - • Editorial quality of added content *336*
  - – Security versus accessibility *337*

7.02 A closer look *338*
(i) What are "accessible" or "alternate" formats? *339*
– The laws are ambiguous *339*
• Chafee amendment *339*
• Instructional Materials Accessibility Act meaning of "print-disabled" *339*
– Sections 504 and 508 of the Rehabilitation Act cover all disabilities *340*
• Different requirements for different disabilities *340*
• Formal qualitative standards for accommodations are not in place *340*
– What can we conclude about accessible alternate formats? *341*
• The IMAA's proposed "national standard format" for school materials *341*
• Needs of Braille makers not likely to go away *342*
• No standards for non-K–12 publishers *342*
– What about Adobe's PDF? *342*
• Protected text makes PDF not accessible *343*
• Section 508 has driven Adobe to address accessibility needs *343*
• Secure PDF remains a problem *343*
– What alternative formats are actually being used today? *344*
• Text formats *344*
• Unstructured ASCII is most common for print works *344*
• HTML coming on strong for electronic works *344*
• W3C/WAI guidelines becoming a *de facto* standard *345*
(ii) So where are we now? *345*
– Many assistive technologies; all require digital content *345*
– Scanner generated ASCII is the dominant file form *345*
– Braille users are unlikely to be moved *346*
– Publishers can anticipate increased demand for electronic files *346*
(iii) What is involved in making a digital work fully accessible? *346*
– Accessibility is a matter of degree *346*
– Steps to making a document accessible *347*
– Degree of accessibility desired drives costs *347*
(iv) Some key problems in dealing with accessibility today *348*
– Much of the Internet is not accessible *348*
– Braille infrastructure is not responsive enough *348*

– Digital Rights Management is in conflict with accessibility needs 348
– Legacy and backlist works can be expensive to convert 349
– Embedded rights 349
– User unfamiliarity 350
– Lack of business models 350
– Chafee confusion 350
(v) What are current drivers for change? 351
– Stronger laws 351
– The Internet 351
– XML technologies 352
– Digital Rights Management 352
– High-stakes testing 352
(vi) Strategic considerations relating to alternative formats 353
– Alternate formats as products for sale 353
– Backlist conversions 354
– Competitive considerations 354
(vii) Strategic perspectives 355
– Minimalist approach 355
– Pragmatic approach 356
– Leadership approach 356
– Underlying considerations 357
• Is outsourcing an option? 357
• Technical resources 357
• Editorial resources 357
• Legacy vs. current work 358
• Halo effect 358
• Business models 359
• User support 359
• DRM support 359
(viii) Summary suggestions 359
– Get organized about accessibility 359
– Get organized about content 360
– Get organized about technology 361
– Get organized about the market 361
7.03 Closing summary 361
7.04 Resources and documents 363
(i) Legislation—federal 363
(ii) Legislation—states 364
(iii) Standards 365

(iv) Making accessible documents *366*
(v) Other resources *366*

Chapter 8. Digital Printing *369*

George Alexander, Executive Editor, the Seybold Report

The technology that is used to print books digitally has been evolving rapidly. Initially, sheetfed monochrome laser printers were the only technology available. But today, high-speed roll-fed devices are often used, full color is becoming more economical, and other printing technologies (such as inkjet) are coming into use. This chapter describes the available printing and binding options and some of the new publishing options they create.

8.01 Overview *369*
8.02 Digital printing technologies *369*
(i) Input: scanning, PostScript, PDF, and other options *370*
– Scanning *370*
– PostScript and PDF *371*
(ii) Imaging technologies for digital printing *372*
– Laser printers and other toner-based printers *372*
– Inkjet *373*
– Elcography *374*
(iii) Other factors affecting choices in digital printing *374*
– Cost *374*
– Speed *375*
– Flexibility *375*
– Appearance *376*
– Binding *376*
8.03 Uses of digital printing *377*
(i) Digital printing vs. offset printing *377*
– Quality *377*
– Speed *377*
– Flexibility *377*
– Cost *378*
– Operational factors *378*
(ii) Printing on demand *378*
– Books on demand: making it work *379*
(iii) Short-run printing *379*
– Applications for short-run printing *380*
• Course packs and custom textbooks *380*
• Advance copies and review copies *380*

• Backlist books 380
• Small publishers and self-publishers 380
– The success of short-run digital printing 381
(iv) Personalized and custom documents 381
(v) The role of the Web 382
8.04 Available printing systems 382
(i) Sheetfed printers 382
– Monochrome sheetfed printing 382
• IBM 3800 383
• Xerox 9700 383
• PostScript and DocuTech 384
• The market today 384
– Color sheetfed printing 385
(ii) Roll-fed digital printing 386
– Monochrome machines 387
• IBM 387
• Oce 387
• Xerox 387
• Nipson and magnetographic printing 388
• Delphax and electron-beam imaging 388
• Scitex Digital Printing and Versamark 388
• Elcorsy Elco 400 389
– Color machines 389
(iii) Cluster approaches 390
8.05 Sales channels for digital book printing 390
(i) Book and journal manufacturers 391
(ii) Book wholesalers and distributors 391
(iii) Vanity presses and self-publishing services 391
(iv) Copy shops 392
(v) Digital printing in the bookstore 392

Chapter 9. Multimedia Publishing 393
Florian Brody, President and CEO, Brody Inc.

Since the format of the book evolved from the continuous scroll to the page-oriented folio, no change in the practice of publishing texts has had such an impact on the way we perceive and use a book as electronic publishing. This paradigm shift changes the way text is perceived in time and space and the integration of text, video, and audio into a multimedia product is a logical step in an electronic medium. Yet it is not the technology that undergoes the biggest change, but the role of the publisher, who has to re-emerge as the *agent of a new medium, still in statu nascendi. In the first phase of multimedia, everybody seemed to be empowered by the new tools and technologies to become a multimedia producer. Most multimedia publications do not live up to the promise of an interactive and integrated experience, but remain an exploration into technologies without a clear goal. It is the publisher who needs to act as the integrator of multiple media types, multiple experts, and multiple industries in order to do his job—to turn an idea into a product and make it public. This chapter gives an overview of the different technologies, standards, and business issues to be considered when extending electronic publishing into multimedia.*

9.01 Overview *393*
9.02 What is multimedia? *394*
- (i) A brief definition *394*

9.03 Deciding on multimedia *395*
- (i) Decision criteria *395*
  - – Will the addition of sound and/or video contribute to the communication of the core message? *396*
  - – Can the target audience be expected to have the necessary hardware and software? *397*
  - – Are the production and/or licensing costs within the budgetary limitations and the production time frame? *397*
- (ii) Expanding the medium—the benefit for the publisher *398*
- (iii) Some background on multimedia *398*
  - – A pioneering product: Beethoven's 9th Symphony *399*
- (iv) On-line multimedia *399*
- (v) Multimedia in digital publishing today *400*
- (vi) Less is more: the power of text *400*
- (vii) Risks and threats of multimedia *401*

9.04 Multimedia experience *401*
9.05 The business of multimedia *402*
9.06 Multimedia technology *402*
- (i) Delivery platforms *403*
- (ii) Multimedia types *404*
- (iii) Media formats *404*
  - – Video formats *405*
  - – Integration of video *405*
  - – Audio formats *407*
  - – Media players *408*
    - • Windows Media *408*
    - • Real Audio/Video *408*

• QuickTime *408*
• MP3 *408*
(iv) Authoring tools and their multimedia capabilities *408*
(v) Multimedia in e-books *409*
(vi) Multimedia authoring software *409*
– Macromedia Flash and Director Shockwave *410*
(vii) Integration and user control *411*
– Control conventions *412*
(viii) Hyperlinks and nonlinear texts *412*
(ix) Standardization *412*
9.07 Rights issues *414*
(i) Usage rights model *415*
(ii) Integration of multiple media means multiple industries *415*
(iii) Digital rights management *415*
– EULA—the End-User License Agreement *415*
– Production issues *416*
9.08 Conclusion *416*

Chapter 10. Content Management & Web Publishing *418*
Bill Trippe, President, New Millennium Publishing
Mark Walter, Consultant, Seybold Consulting Group

This chapter tackles two subjects—the process of publishing on the Web, and the technology called content management that has emerged to address Web publishing. Publishers who have moved to the Web have found it brings all the challenges of print publishing—and some new challenges that are unique to the Web. Content management technology has grown to meet these challenges, but has also brought with it technical complexity that is new for some publishers. This chapter outlines the technical issues regarding content management, and gives the readers a framework for understanding how content management technology can help their publishing processes.

10.01 Introduction to content management *418*
(i) What is content management? *418*
(ii) Why content management systems are needed *419*
– Increasing complexity while maintaining consistency *420*
• Lack of control over precise rendering *420*
• Linking *421*
• A different reading experience *421*
• Navigation *421*
• Personalization *422*
• Information architecture *422*

(ii) Proprietary CMS development issues *447*
- Application service providers (ASP) *448*

(iii) "Your mileage may vary"—opting for open source *449*
- Advantages of an open source CMS *450*
- Disadvantages of an open source CMS *450*
- Cost *451*
- Control *451*
- Support *451*

(iv) Commercial content management systems: additional factors for consideration *452*
- Price, Size, Scale *452*
  - Upper-tier content management systems *453*
  - Mid-market content management systems *453*
  - Low-priced content mangement systems *453*
- Technology *453*
- Industry specialization *453*

10.06 Post-implementation issues *454*

Chapter 11. Electronic Books & the Open eBook Publication Structure *455*

Allen Renear, University of Illinois, Urbana-Champaign
Dorothea Salo, University of Wisconsin, Madison

Electronic books, or e-books, will soon be a major part of electronic publishing. This chapter introduces the notion of electronic books, reviewing their history, the advantages they promise, and the difficulties in predicting the pace and nature of e-book development and adoption. It then analyzes some of the critical problems facing both individual publishers and the industry as a whole, drawing on our current understanding of fundamental principles and best practice in information processing and publishing. In the context of this analysis the Open eBook Forum Publication Structure, a widely used XML-based content format, is presented as a foundation for high-performance electronic publishing.

11.01 Introduction *455*
11.02 OEBPS in a nutshell *457*

(i) Basic description of OEBPS *457*
- What it is *457*
- What it does *458*
- Who developed it *458*
- What it is for *458*

• Interactivity 422
• Other media 423
• Cobranding and syndication 423
– Volume of content and contributors 424
– Static vs. dynamic pages 425
• CGI, ASP, and JSP 426
• Static or dynamic? Both 426
• The three phases of content management 426
• Personalization 427
10.02 Types of Content Management Systems 427
(i) Web-centric content management systems 428
– Departmental 429
– Midrange systems 429
– Enterprise-level CMSs 429
(ii) XML-based content management systems 430
(iii) Genre-specific content management systems 431
– Newspaper systems 431
– Magazine systems 432
– Reference publishing systems 432
10.03 Benefits of content management systems 433
(i) Organizing files 433
(ii) Separate form from content 434
(iii) Automate mundane tasks 435
(iv) Making it easier to increase sophistication 437
10.04 Issues to consider in content management 438
(i) Workflow 438
(ii) Views of the repository 440
(iii) Version control and library services 441
(iv) Scheduling: publication and kill dates 442
(v) Templates and stylesheets 442
(vi) Text-authoring interfaces 443
– HTML forms 444
– HTML editors 444
– Word processors 444
– XML editors 445
(vii) Personalization 445
(viii) Syndication 446
10.05 Evaluating a content mangement system 447
(i) Factors to consider in evaluating content management systems 447

– Why we need yet another specification *458*
– What it is, in a little more detail *459*
(ii) Some very important facts about OEBPS *459*
– Specialized XML vocabularies are encouraged *459*
– XHTML gets special consideration *460*
– OEBPS is used in several different ways *460*
– Processing conformance targets *reading systems* not *reading devices* *460*
– OEBPS does not *necessarily* compete with Adobe's PDF *461*
– Multimedia and active content is allowed, with conditions *461*
– OEBPS Publications can be easily created from XML documents *462*
– OEBPS 1.2 substantially improves formatting capability *463*
– OEBPS is widely used in the publishing Industry *463*
– OEBPS 2.0 will bring further improvements *463*
11.03 Electronic books in general *464*
(i) E-books: Brief history and future of the idea *464*
(ii) Terminology and scope: What is an e-book? *465*
(iii) Advantages of electronic books *466*
– Capacity *466*
– Manufacturing *466*
– Distribution *466*
– Cost *467*
– Intelligent viewing *467*
– Intelligent navigation *467*
– Hypertext *467*
– Retrieval *467*
– Currency *468*
– Multimedia *468*
– Interactivity and special processing *468*
– Accessibility *469*
11.04 Thinking clearly about e-books *469*
(i) Navigating the hype *469*
– The replacement red herring *470*
– The bed-beach-bath red herring *470*
– The lots-of-advantages red herring *470*
(ii) Why we know that electronic books will happen *471*
– Hardware improvements *471*
– Software improvements *471*
– Ubiquity of reading devices *472*

– Critical mass of content 472
– Data standards and interoperable tools 472
– Culture of electronic reading 472
(iii) What we don't know about *how* they will happen 473
– Form factor 473
– Publishing markets 473
– Industry structure 473
– Business models, intellectual property, digital rights management, and security 474
11.05 The format problem 474
(i) Introduction 474
(ii) Categories of formats 475
– Logical vs. presentational 475
• The logical approach separates structure from presentation 475
• Descriptive markup implements the logical approach 476
• SGML and XML define descriptive markup languages 476
• The logical approach yields many advantages for publishing 476
• Presentation formats are still important 476
• Compromise strategies may be necessary 476
– Revisability and representation 477
• Revisable vs. nonrevisable formats 477
• Text vs. binary formats 477
(iii) Formats up- and downstream in e-book production 478
– Formats found upstream in the e-book production process and upstream in the print production process 479
• Revisable binary formats and their derived revisable text formats 479
• Revisable text formats (general purpose) 480
– Formats found upstream in the e-book production process and downstream in the print production process 482
• Revisable binary formats and their derived revisable text formats 482
• Nonrevisable text and binary formats 483
– Formats found downstream in the e-book production process 483
• Nonrevisable binary formats 484
• Revisable text formats 484

(iv) Format-related challenges *484*
– Functionality, innovation, and competitive differentiation vs. interoperability *484*
– Functionality vs. current reality *485*
– The problems of format conversion *486*
(v) Solving the format problem *487*
11.06 The OeBF Open eBook Publication Structure *488*
(i) History *488*
– The development of OEBPS 1.0 *488*
– The formation of the OeBF *489*
– Formation of the Publication Structure Working Group (PSWG) *489*
• Officers *489*
• Composition and current activities *490*
– History of OEBPS releases and adoption *490*
(ii) Purpose and nature of OEBPS *491*
– Official Purpose and Scope *491*
– Empirically identified objectives *491*
• Support consumer confidence in performance and interoperability of devices and books *492*
• Create a critical mass of content *492*
• Limit burden on content providers *492*
• Support content provider need for reliable high-quality performance *493*
• Limit burden on developers of reading software *493*
• Support distinctive new functionality *493*
• Maintain equitable opportunities for competitive differentiation *493*
• Position industry practices to evolve with emerging standards *494*
• Provide an aesthetically satisfying reading experience *494*
• Support other languages and writing systems *494*
• Support access by readers with disabilities *494*
• Have an *immediate* and *direct* impact on the creation oflourishing e-book industry *495*
• An overarching secondary goal: reconcile conflicting primary goals *495*
– General nature of OEBPS: an application of already existing standards *495*
– Summary: the purpose and nature of OEBPS *496*

(iii) Terminology *497*
- OEBPS Publication *497*
- OEBPS Package *498*
- OEBPS Document *498*
- Basic OEBPS Document *498*
- Extended OEBPS Document *499*
- OEBPS Core Media Type *499*
- Reading System *499*
- Reading Device *500*
- Content Provider *500*
- Reader *500*
- Deprecated *501*

(iv) The OEBPS processing model *501*
(v) The component standards *501*
- XML *501*
- XHTML/Basic OEBPS Document Vocabulary *503*
- CSS/OEBPS style language *504*
- XML namespaces *504*
- MIME media types *505*
- Unicode, UTF-8, UTF-16 *506*
- Dublin Core *506*

(vi) The package *507*
- Components of the Package *507*
  • Package identity *507*
  • Metadata *507*
  • Manifest *508*
  • Spine *508*
  • Tours *508*
  • Guide *508*
- Example of a Package file *509*

(vii) Toward OEBPS 2.0 *511*
- Metadata modularity *511*
- Internationalization *512*
- Inter- and intra-publication linking *512*
- Navigation *513*
- Packaging and compatibility *513*

11.07 In conclusion *513*
11.08 Some advice for e-book publishers *514*
(i) What you must know *514*
(ii) What you must do *515*

11.09 For more information *517*
(i) The Open eBook Forum *517*
(ii) The Open eBook Publication Structure *517*
(iii) Current research and analysis: functionality *518*
(iv) Current research and analysis: users *518*
(v) SGML/XML and the logical approach *518*
(vi) Book-related SGML/XML vocabularies *519*
(vii) Classics *519*
(viii) Wider issues *520*

Chapter 12. Archiving *521*
Heather Malloy, Digital Archive Manager, John Wiley & Sons

Maintaining ownership of commercially viable digital assets is increasingly important. As digital files have become the de facto standard for use and reuse of products published both in print and electronic formats, libraries and publishers are working to create a viable way to store and preserve digital assets. Publishers have additional reasons for creating the process and infrastructure for storing their digital materials: the potential for reduced operating costs and increased profit margins associated with reusable content and the increased revenue derived from licensing, selling, or otherwise making available content that is centrally available. This chapter will focus on the issues facing publishers, but will also provide an overview of the wider issues involving archiving.

12.01 The importance of archiving *521*
(i) Commercial needs *521*
– Ownership of assets *522*
– Reuse *522*
• Direct revenue *522*
• Indirect revenue *523*
• Future use *523*
(ii) Long-term preservation and access *524*
12.02 Other concerns for archiving *524*
(i) Are commercial needs and scholarly preservation mutually exclusive? *524*
(ii) Publishers and libraries cooperating *526*
(iii) Should everything be archived? *526*
12.03 Where to implement the archive *527*
(i) Third-party vendors *527*
(ii) Noncommercial repositories *528*

(iii) In-house digital archive *528*
(iv) Combining options *529*
12.04 Technology issues *529*
(i) The problem of obsolescence *529*
– Migration *530*
– Emulation *530*
(ii) Format vs. content *530*
(iii) Formats *531*
– Sample of format types *531*
• PDF *531*
• XML *532*
• Page layout and composition programs *533*
• Graphics software *533*
• Other software and applications *534*
(iv) Storage media *534*
(v) Metadata *535*
– Overview of some relevant metadata and archiving standards *536*
• ONIX (On-line Information Exchange) *536*
• Open Archival Information System reference model *536*
• Dublin Core Metadata Initiative *537*
• Open Archives Initiative protocol for metadata harvesting *537*
12.05 Issues in development and implementation *537*
(i) Cost *537*
(ii) Quality control *538*
(iii) Workflow *539*
– Publishers *539*
– Service providers to publishers *540*
– Customers *541*
– Authors *541*
(iv) Content management *542*
(v) Copyright and DRM *542*
12.06 Conclusion *542*
12.07 Resources *543*
(i) On-line References *543*
(ii) Organizations and Committees *543*

Chapter 13. The Legal Framework: Copyright & Trademark *546*
William S. Strong, Partner, Kotin, Crabtree & Strong, LLP, Boston, Massachusetts; Author, The Copyright Book: A Practical Guide

The purpose of this chapter is to give a succinct overview of the various legal doctrines that apply to digital publishing and to make the concepts and basic rules as accessible as possible. This chapter does not substitute for legal advice on any specific matter that the reader may encounter; rather, it is intended to help the reader be more alert to legal issues that may arise, and be a more informed consumer of legal services. It does not, of course, purport to address all the legal issues that a publisher will encounter. Publishers are businesses, and must deal with all the laws that affect businesses of every kind. Nor does this chapter address some issues that are common to everyone who does business on the Internet, such as the laws of consumer privacy, e-commerce, and the like. The main focus is on copyright, as that creates the property that is the currency of publishing. This chapter will also deal with other intellectual property principles and with some other concerns, such as libel, that are uniquely central to publishing.

13.01 Copyright *546*
- (i) Overview *546*
  - – What can be copyrighted *547*
  - – Copyright term *547*
  - – Copyright is national and international *547*
- (ii) Criteria for copyrightability *548*
  - – Expression *548*
  - – Originality *548*
  - – Creativity *548*
  - – Variety of possible expressions *549*
  - – Fixed in a medium *549*
  - – Nontrivial quantity *549*
  - – More than purely functional *549*
  - – U.S. government works *550*
  - – Copyright is automatic *550*
    - • Copyright notice *550*
    - • Registration *550*
- (iii) Who is the copyright owner? *551*
  - – Works for hire *551*
    - • Works by regular employees *551*
    - • Works by surrogate employees *551*
    - • Commissioned works *552*
- (iv) Copyright registration *553*
- (v) The rights of a copyright owner *553*
  - – Copies *553*
  - – Derivative works *553*

– Distribution 554
– Public performance 554
– Public display 554
– Importation 555
(vi) Dealing in rights: Assignments and licenses 555
– The importance of getting it clear, precise, and in writing 556
• Transfer or sublicensing 556
• The *Tasini* case 556
• The *Rosetta Books* case 558
• The impact of *Tasini* and *Rosetta* 558
– Language-specific licenses 558
– A few thoughts on licensing strategies 559
• Should the author or publisher retain electronic rights? 559
• What about foreign language e-rights? 559
• What if the book is intended as an e-book? 559
• Scholarly journals should keep electronic and print rights together 560
• Personal posting of journal articles 560
• World rights in foreign languages 560
• Electronic rights in foreign languages 560
– Nonexclusive licenses 561
– Permissions 561
(vii) Fair use 562
– Fair use in the digital era 563
• Criticism 564
• Entertainment 564
• Illustration 564
• Appropriate quantity and quality 565
• Promotional uses 565
• Other uses 566
(viii) Protecting copyrights against infringement 567
– Copy prevention and encryption 567
• Nonfunctional protection mechanisms 568
• Other exceptions 568
• Fair use in the context of encryption 568
– Closing infringing Web sites 568
– Protecting copyright management information 569
– Hacking 569

13.02 Trademark law 570
- (i) The strength of the mark 570
  - Book titles 571
  - Business names and URLs do not guarantee trademarks 572
  - The issue of public confusion 572
- (ii) Avoiding trademark conflicts 572
  - Not all trademark searches are the same 573
    - Commercial searches 573
    - Internet searches 573
    - Searching official registries 573
    - International searches 574
- (iii) Establishing the mark 574
  - Use, "intent to use," and registration 574
  - How to register a trademark 575
    - Protecting book titles on the Web 575
  - Trademark notice 575
    - Using and acknowledging others' trademarks 576
  - Use of trademarks in metadata 576

13.03 Other laws 577
- (i) Libel 577
  - Avoiding libel 578
    - Linking 578
- (ii) Rights of privacy and publicity 578
  - Publicity 579
  - Foreign laws 579
    - Moral rights 579
    - Other foreign laws 581
    - Avoiding problems with foreign laws 581
    - The Yahoo case 581

13.04 Lawsuits: Is there nationwide jurisdiction? 582

13.05 Contracting with customers 582
- (i) Licensing the individual reader 582
  - Keeping copies 583
  - Printing by the user 583
  - Using extracts 583
  - Sharing or reselling 584
- (ii) Licensing the user organization 585
  - Interlibrary loans 585
  - Reserve copies 586
  - Other uses 586

13.06 Conclusion 586

Chapter 14. International Issues *587*

Robert E. Baensch, Director, Center for Publishing, New York University

This chapter highlights the twelve key elements that are most important for international communications, online publishing, and e-commerce on the Internet. The size of organization or location of country is irrelevant, because the World Wide Web is a stateless network and framework that goes beyond the physical location of electronic resources and information by connecting millions of computers into a seamless global network. Statistics on Internet users worldwide, with a focus on Europe, Asia, and South America, establish a meaningful framework for what it really means to publish globally. A review of the international publishing activities of professional, legal, scientific, technical, and medical journal publishers offers realistic working examples for publishers who are considering the development of their international on-line business. A concise analysis of geographic, cultural, language, economic, technological, and legal factors provides perspective on the global environment for digital publishing.

14.01 Overview *587*
14.02 Internet users worldwide *588*
- (i) Europe *591*
- (ii) Asia *591*
- (iii) South America *592*

14.03 The STM industry leaders *594*
14.04 Establishing the Web Site *596*
- (i) Information distribution *597*
- (ii) an interactive communications channel *597*
- (iii) New sales and distribution channels *598*
- (iv) E-commerce *599*
- (v) Market research *599*

14.05 Understanding the global environment *600*
14.06 Geographic and country priorities *601*
14.07 English and other languages *602*
- (i) Foreign language use in Europe *603*
- (ii) Other barriers to global use *604*

14.08 New economics of information services *605*
- (i) Consolidation of publishers and their customers *606*

14.09 Worldwide on-line advertising *606*
14.10 Marketing on the Internet *608*
- (i) Brazil *608*
- (ii) China *608*
- (iii) Japan *609*

(iv) Poland 609
(v) South Africa 610
(vi) South Korea 610
14.11 International information sources 610
(i) Financial and economic information for international markets 611
(ii) Regional or country information 611
(iii) International business resources 612
14.12 Internet publishing law 612
(i) Copyright 612
(ii) Trademark 613
14.13 Conclusion 614

Chapter 15. Digital Rights Management *616*
Paul Hilts, Former Technology Editor, Publisher's Weekly

One of the promises of the digital revolution in publishing is that it has the potential to give owners of content—authors, publishers, aggregators—new, robust tools for more intelligent, efficient, and effective management of their content. Nowhere have the benefits of new technology been more eagerly awaited than in the area of rights management. However, content owners have been groping not only for the right technology but for the right applications of competing technologies—as well as dealing with changing business models, customer resistance, legal challenges, and implementation questions. This chapter explores the fundamentals of Digital Rights Management and the different DRM technologies available in order to help publishers and their partners implement the right technology in the right manner.

15.01 Overview: What is DRM? 616
(i) Controlling access to content 617
- Taking payment 618
- Identifying and authenticating users 619
- Specifying valid uses 619
- Measuring use 620
- Deterring piracy 620
(ii) Approximating physical distribution models 620
- Disabling infinite perfect copies 620
- Lending, read-only, etc. 621
(iii) Inventing new distribution models 622
- Dynamic pricing 622
- Demographics-per-view 623

– Post-PC devices (e-books) *623*
– Business-to-business (syndication) *623*
15.02 Rights-based business models *624*
(i) Pre-digital rights management *624*
– Rights management in physical media *625*
• Physical copies of books are imperfect and expensive *625*
• Limiting "infringement technology" *626*
– Contractual rights *626*
– Compensation models *627*
– Government intervention: compulsory licensing *628*
– Unmanaged rights: copyright law *628*
(ii) DRM-enabled business models *629*
– Paid downloads *629*
– Subscriptions *630*
– Pay-per-view/listen *631*
– Usage metering *631*
– Peer-to-peer/superdistribution *632*
– Rights licensing *633*
15.03 DRM technology *633*
(i) Reference architecture *634*
– The content server *634*
• Product information: metadata *635*
• DRM packager *635*
– The license server *635*
– The client *636*
(ii) DRM functions: identification, encryption *637*
– Identities *637*
• User authentication *638*
• Device authentication *639*
– Measuring use for market intelligence *640*
– Encryption and DRM *641*
• How good is encryption? You do the math *641*
• Not good enough *642*
• Good enough *643*
– Watermarking vs. encryption *644*
– Watermarking plus encryption *644*
15.04 DRM standards *644*
(i) XrML: the eXtensible rights Markup Language *645*
(ii) DOI: the Digital Object Identifier *646*
(iii) ICE: Information and Content Exchange *648*

(iv) Other DRM standards *650*
- ODRL: Open Digital Rights Language *650*
- PRISM: Publishing Requirements for Industry Standard Metadata *650*
- <indecs>2RDD *651*
- W3C *651*

15.05 Legal developments: important legislation *652*
(i) Milestones *652*
- 1996: WIPO copyright treaties *653*
- 1998: Digital Millennium Copyright Act (DMCA) *653*
- 2000: Electronic Signatures Act *654*
- 2001: European Copyright Directive (ECD) *654*
- 2002 (proposed): Consumer Broadband and Digital Television Promotion Act (CBDTPA) *655*

(ii) Legislating DRM? *655*
(iii) Current cases *656*
- New York Times v. Tasini *656*
  - Effects *656*
- Random House v. RosettaBooks *657*

15.06 Vendors: DRM solutions for publishers *658*
(i) Adobe *658*
(ii) Microsoft *659*
- Third generation: "Unified DRM" *660*

(iii) Notes on other systems and vendors *661*
- InterTrust *661*
- IBM: EMMS *661*
- SealedMedia *661*
- OverDrive Systems *662*
- Savantech *662*

15.07 The state of the market *662*
(i) Likely market development scenario *664*

15.08 DRM Implementation Issues *664*
(i) Is encryption necessary? *664*
(ii) Consider watermarking *665*
(iii) Seeking a vendor: Validation criteria *665*

Glossary *667*
Bibliography *709*
Index *719*

Credits

Columbia University Press
William B. Strachan *Publisher*
Linda Secondari *Creative Director, Manufacturing and Technology*
Stephen H. Sterns *Editor, Electronic Reference*
Richard Hendel *Designer*
Paul Berk *Copyeditor*

Impressions Book and Journal Services
William E. Kasdorf *President*
David Nelson *Electronic Publishing Manager*
Christie Pinter *Project Coordinator*
Chris Cott *Systems Specialist/Senior Pagination Specialist*
Sharon Hughes *Indexer*

Open Book Systems
Laura Fillmore *President*
Dee Landergren *Vice President of Systems*
Sunny Gleason *Manager of Programming*
David G. Cassidy *Project Lead*
Mircea Baciu *Programmer*
Danielle Hull *Information Designer*
Kendall Dawson *HTML programmer*

Contributors

George Alexander

Alexander is executive editor of *The Seybold Report* and has been involved with publishing technology since 1977. He began his career in publishing as systems manager for a daily newspaper and later served as director of publishing systems for Harper & Row. During his twenty years with the Seybold organization, he has written articles and done consulting projects on many areas of publishing technology. His particular areas of interest include digital printing, computer-to-plate systems, color management, cross-media publishing, and systems to support large-scale publishing. His consulting clients have included many well-known publishing organizations, including the *New England Journal of Medicine,* World Book, Scholastic, and HarperCollins.

Robert Baensch

Baensch is Associate Professor of Publishing and the Director of the Center for Publishing, New York University. Before joining NYU, he was a publishing consultant and president of Baensch International Group Ltd. New York. He was senior vice president for marketing for Rizzoli International Publications, Inc. Prior to that as director for publishing at the American Institute of Physics from 1988 to 1991, he was responsible for over 60 journals, a book program and database information services. From 1983 to 1988, Baensch was vice president-marketing of Macmillan Publishing Company where, in addition to a full range of marketing and sales responsibilities, he directed the Macmillan Software Company and English as a Second Language Multimedia Program. Before 1983, he was president of Springer Verlag New York, and from 1968 to 1980 vice president and director of the International Division of Harper & Row Publishers, Inc. Baensch started his publishing career with the McGraw-Hill Book Company, where he was manager of the Translation Rights Department and editorial director of the International Division. He has served on the Board of Directors of the Association of American Publishers and the Society of Schol-

8 Digital Printing

GEORGE ALEXANDER
Executive Editor, the Seybold Report

The technology that is used to print books digitally has been evolving rapidly. Initially, sheetfed monochrome laser printers were the only technology available. But today, high-speed roll-fed devices are often used, full color is becoming more economical, and other printing technologies (such as inkjet) are coming into use. This chapter describes the available printing and binding options and some of the new publishing options they create.

8.1 OVERVIEW

The digital printing of books is a tantalizing concept. One can imagine walking into a neighborhood bookstore (or perhaps a neighborhood copy shop) and ordering virtually any book ever published. Within five minutes, the book could printed, bound, and in your hands. For publishers, there is the possibility of never having overstocks to dispose of. For educators, the technology can provide custom books for each class, or even for each student.

The technology to support these possibilities is not completely available yet (or if it is, the costs are still prohibitive for many applications). But the technology is getting very close, and costs are coming down rapidly. In this chapter, we will look at the digital printing technologies that are available today and some that are on the way. The associated input techniques and binding technologies are also discussed.

Increasingly, digital printing is taking work away from the traditional offset process. We will discuss the advantages and disadvantages of both approaches and list some of the applications for which digital printing is clearly superior. The shift from offset to digital printing also means changes in the where books are produced and how they are distributed. These, too, will be touched on.

8.2 DIGITAL PRINTING TECHNOLOGIES

The technology of digital printing has been evolving rapidly. It is not just the printing techniques themselves that are in flux, but also the associated technologies (such as those involved in scanning and binding). This section describes the technologies behind digital printing and some of the important related issues.

(I) INPUT: SCANNING, POSTSCRIPT, PDF, AND OTHER OPTIONS

Various methods and technologies are available to prepare material for digital printing. The most fundamental fact, however, is that digital printing is done directly from digital files. Content is either created in a digital format or it must somehow be converted from print to a digital format.

For the publishing industry, most files created in a digital format end up either in the ***PostScript*** language, or in its close cousin, the Portable Document Format (***PDF***). Both languages are the invention of Adobe Systems. Adobe, although it makes the specifications of these languages public, keeps tight control over the direction and details of both. Content that does not originate in digital form—that is, printed originals—must be digitized. Most often, this is done simply by scanning, which results in a bitmapped file, usually a TIFF.

The chapters on **data conversion** [5.3.1.1] and on **composition and graphics** [6.11.1] discuss the details of the processes of scanning and conversion by rekeying; for conversions, this section focuses on digital conversion rather than manual rekeying.

1. Scanning

Scanning—that is, making an image of a printed page rather than using ***OCR*** software to create a digital text file—is generally easy to accomplish, but it has a significant drawback. No matter how good the scanner, and how good the page being scanned, the image created via scanning is never quite as good as the same image created directly from the source data. Some degradation is inevitable in the image-then-scan sequence. For example, the pages are usually not precisely straight, since scanning is a mechanical process; the resulting images usually need ***deskewing***. Likewise, it is virtually impossible not to pick up specks of dust and other flaws in the image; this requires ***despeckling***. Finally, the pages usually need to be aligned into the position required by the new product, with uniform margins and with right- and left-hand pages being differentiated. In spite of great improvements in image-processing technology, these issues are still true. For that reason, scanning is now rarely used as an input method for digital printing. The big exception to this rule is reprinting books (and other documents) where no digital original is available—and this is the case for most reprinted books today.

The scanning process usually produces a bitmapped file known as a ***TIFF*** file. One-bit TIFFs are often used for text; they capture the image simply in terms of black or white pixels. Eight-bit TIFFs are more sophisticated: they are able to capture ***grayscale*** information (shades of gray) and thus are more appropriate when photographs appear in the originals. However, when text is scanned as 8-bit, it is often a bit blurry. The best results are accomplished by scanning at 8-bit and then downsampling the text areas to 1-bit (ideally, smoothing out the jagged edges of the type and despeckling the surrounding white areas) and leaving the image areas as 8-bit data.

2. PostScript and PDF

PostScript is a very general-purpose page-description language, and its incorporation into the first Apple LaserWriter was an important factor in the development of "desktop publishing" in the 1980s. The versatility of PostScript was important in its early success—it was much more flexible than the other languages used to drive printers in those days—but its very flexibility made it somewhat unpredictable. ***PDF***, which can be thought of as an efficient and predictable subset of PostScript, is gradually taking over PostScript's role as the normal file format for output to a digital printer.

Although other languages are used to drive office printers (notably Hewlett-Packard's ***PCL***) and to drive high-speed printers in data centers, PostScript and PDF are the norm for publishing applications. That doesn't mean, however, that all publishers provide their printers with PostScript or PDF files. Most are still providing native files from their layout package (e.g. QuarkXpress, Adobe InDesign, or Microsoft Publisher) and letting the printer produce the final PostScript or PDF file. This gives the printer more flexibility in fixing errors and avoiding production problems. New dialects of PDF (such as PDF/X-1a and PDF/3), along with the appropriate software support, promise to make it easier for publishers to supply files that will be suitable for printing. The chapter on **composition and graphics** [6.12.11.5] discusses these technologies in more detail.

It should also be noted that it is possible for PostScript or PDF files to consist mainly or only of bitmapped data (usually TIFFs), since the ability to incorporate TIFF images is an important feature of both of these technologies. This can be the cause of great confusion. Generally, when PostScript or PDF files are being referred to, it is assumed that they are text-based files that were generated from a digital page-makeup process; such files are editable and searchable, and use fonts and vector data to conform to the resolution of whatever output device is being used. If, on the other hand, they are simply scanned data captured as TIFFs and then enclosed in PostScript or PDF "wrappers," the data they contain is not editable or searchable and will not adapt; their resolution is dependent on how much data was captured in those original scans. Especially when PDF files are being provided to or received from a new source, it is important for all parties to be aware whether the underlying data is true text-and-vector data or bitmapped image data, or both.

(ii) IMAGING TECHNOLOGIES FOR DIGITAL PRINTING

The majority of digital printing is done with toner-based devices, but a growing amount is being accomplished in other ways. Each approach, and there are many, has a specific set of advantages and disadvantages.

1. Laser printers and other toner-based printers

The most common kind of toner-based printing device is the laser printer. A laser beam bounces off a spinning mirror and sweeps across a charged photoconductive drum or belt, which is discharged wherever the laser light touches it. Charged toner particles adhere to the drum or belt where it was exposed. The toner is then transferred to the paper (sometimes via an intermediate drum or belt). Once on the paper, the toner is fused (melted in place), usually by a hot roller. The most common production-speed monochrome laser printers are made by Xerox, Canon, and Ricoh. (Specific printers are discussed in the section on **available printing systems** [8.4] below.)

A color machine works on the same principles as a monochrome one, except that there are four different toners (cyan, yellow, magenta, and black). There are often four different lasers and four different drums (or belts) as well, but sometimes these components are used in common for the four toners.

The same sort of device can be built with an array of light-emitting diodes (***LED***s) instead of a laser and a spinning mirror. Heidelberg and Oce use this technology in their monochrome devices. And it is not even necessary to have a light source or a photosensitive drum. One supplier (Oce) has a color printer that uses a drum whose surface is encircled by fine electrical channels. The channels are rapidly and precisely charged and discharged as the drum turns, to attract charged toner to the appropriate spots. Another supplier (Delphax) uses a special charge-plate with tiny holes to deliver packets of electrons to charge a rotating drum in its electron-beam approach. Yet another (Nipson) uses magnetic toner and a steel drum with a special surface that can be magnetized via a special writing head. This technology is not suitable for color printing, since the toner has to be magnetic and magnetic materials of the appropriate colors are not available.

The quality possible with toner-based printing is limited by the resolution of the printer, which is usually stated in dots per inch (dpi). To reproduce photographs adequately, a toner-based printer must generally have a resolution of at least 600 dpi. This level of resolution is required because toner is a "binary" material—it is either present or not, and if it is present, it is at full strength (solid black). To give the appearance of various levels of gray (in a monochrome device) or shades of color (in a color device), the toner must be applied in "halftone dots" of varying size. If a printer's resolution is much below 600 dpi, the halftone dots are objectionably large. Simliarly, artifacts called "jaggies" are visible in the diagonal edges of text and line art when it is printed on machines with a resolution below 600 dpi. The appearance of both halftones and line art improves detectably as the resolution increases to about 1,000 dpi, but higher resolution than that has little visual effect.

To be useful for book printing, a printer must be resonably fast. A ten-page-per-minute printer could be used to print a 300-page book, but it would take half an hour. In most applications, that would be much too slow. Most book printing is done on machines that print upwards of 100 pages per minute (ppm). Sheetfed printers are available that run up to 180 ppm and roll-fed ones can go well over 1,000 ppm. (Details on the available machines are given **below** [8.4].) The fastest machines are not toner-based, but rather inkjet printers.

Even the fastest of all the digital machines can only match the speed of a moderately slow offset book press. But what digital printing gives up in speed, it compensates for in other advantages, such as flexibility, cost-effectiveness in short runs, and simplified binding processes.

2. Inkjet

Inkjet printing is an entirely different approach to digital printing. Tiny droplets of ink are squirted directly onto the paper from a printhead containing hundreds or thousands of nozzles, each under electronic control.

Most inkjet printing is done using ***drop-on-demand*** technology. Droplets are forced out of each nozzle as they are needed, either by heat (thermal drop-on-demand) or pressure (piezo-electric drop-on-demand). So far, though, the machines using drop-on-demand technology are just not fast enough for most publishing applications, though they are quickly becoming the standard way to produce color proofs. For really fast printing, ***continuous-inkjet*** technology is required. In this technology, all the nozzles in the printhead emit a continuous stream of droplets, but most are electrostaticly deflected back into a gutter and recirculated. Only the undeflected droplets continue on to the paper.

The pioneer in the use of continuous-inkjet systems for publishing is Scitex Digital Printing. Initially, the company supplied partial-page systems for applications like addressing and the printing of variable-text letters, but in recent years they have expanded into full-page systems. Their roll-fed monochrome (Versamark) and color (Business Color Press) systems are not only the first inkjet systems to be used for publishing, they are also the fastest available digital devices of any kind, capable of printing 500 to 1,000 feet per minute—

the equivalent of several thousand pages per minute. Drop-on-demand systems have yet to come near these speeds, though there is much promising development work in this area.

The resolution of the Scitex inkjet machines is just 300 dpi—low, relative to almost any toner-based machine used for book printing. The image quality of the printed result is limited by the resolution. The same will probably be true of other high-speed inkjet machines being readied for the market. But most inkjet technologies have an advantage that toner-based printing does not. Inkjet devices can have variable droplet size (or droplets that consist of variable numbers of sub-droplets) and this allows a certain amount of variation in shades of color or gray without resorting to hafltone dots. By exploiting this "grayscale" capability, inkjet vendors can produce output whose quality (especially where images are concerned) matches that produced by higher-resolution devices. So far, however, no devices that exploit this technique and that would be suitable for book printing are commercially available.

3. Elcography

Finally, there is another unique technology which is not related to either toner-based or inkjet printing. It is called "elcography," and the first elcographic press is being delivered to the first customer as this is written. A distinguishing feature of elcography is that it can print 8-bit greyscale data, reflecting the photographic background of its creators. The technique was invented by the Canadian company Elcorsy, and it is described in more detail in the section on **roll-fed printing** [8.4.2].

(iii) OTHER FACTORS AFFECTING CHOICES IN DIGITAL PRINTING

Although imaging technologies are a key factor in digital printing, other technology issues can play an important role when deciding what system to buy or use. These include costs, speed, flexibility, appearance issues, and binding issues.

1. Cost

Though capital costs are important in digital printing (color devices can range from the hundreds of thousands of dollars into the millions), it is running costs that tend to dominate discussions of the use of digital printing. It is not hard to see why: in many markets, digital printing is competing with offset, and offset's cost per full-color page (once the job is on the press and running) is only a penny or two per letter-size image. Throughout the 1990s, the cost of running a digital press generally exceeded $0.25 per page (and of course the ultimate customer had to be charged substantially more than that if the transaction was to be profitable for the printer). This meant that digital printing was confined to two small niches: very short runs (usually less than 500)—where offset's higher setup costs come into play—and variable-data printing. Now, with the latest generation of digital devices, page costs have come down substantially. The vendors say they will be in the under-ten-cents range for a letter-sized full-color page. If so, digital printing will be competitive at 2,000 copies—not yet the heart of the offset market, but a substantial inroad all the same.

Monochrome digital printing is already competitive with offset in most markets and, in fact, has almost completely eliminated offset in some. Sheetfed monochrome devices—notably the Xerox DocuTech family—displaced much of the low page-count work previously done on offset duplicators; and the high-speed monochrome machines from IBM, Oce, and others are starting to have a similar effect on short-run book printing.

2. Speed

For some applications, speed is important. A speed of 120 ppm sounds impressive, but if you are printing 600-page books on a printer that speed, it can only produce twelve copies per hour. At that rate, a run of 2,500 books would take well over a week, running three shifts and seven days. Only roll-fed printers have the speed necessary for book production in substantial volume. On the other hand, inexpensive sheetfed machines open up interesting new opportunities for very short runs and for printing individual books on demand.

3. Flexibility

Substrate flexibility—what materials can be printed on, and in what form—is an issue with some kinds of digital printing. Not every kind of paper will work in every machine. Some papers do badly with certain toners or certain fusing processes. Some tend to jam in specific machines. Many companies involved in digital printing rely on a qualification process, carried out by their printer and paper suppliers, to make sure that the papers they select will work in their machines. In general, newer machines are more flexible in terms of dealing with surface textures and paper weights than previous models. One supplier, Nipson, has found a niche in monochrome printing on hard-to-image substrates. The combination of its unique magnetographic technology and its low-heat "flash-fusing" approach mean that it can handle a wide variety of plastics and other materials.

Substrate flexibility also means the ability to run a mix of substrates, either in a single job or in successive jobs. Vendors have been adding extra feed trays to their machines, either as standard equipment or as options, in an effort to provide more of this kind of flexibility. This is one area where sheetfed machines have the upper hand. Roll-fed devices cannot mix paper stocks in a job, and it takes several minutes to switch them from one type of paper to another, whereas the same switch is instantaneous for a sheetfed machine with multiple input hoppers.

4. Appearance

The appearance of the printed page can be a very important issue in some applications, and of little significance in others. For a decade, digital-press vendors have been striving to match the appearance of offset-printed pages, since the failure to do so has been a stumbling block in their sales efforts. The current top-of-the-line machines from HP/Indigo, Xerox, and Nexpress do a good job in this respect. Among the problems that have been overcome are: toner sheen (the overly shiny appearance of heavy concentrations of toner), streaking, surface oil (remaining on the paper from the fuser rollers), ***halftone*** quality problems, and ***color gamut*** limitations. Some or all of these appearance problems are still to be found in lower-priced devices. In addition, software issues can contribute to appearance issues such as color problems, image resolution problems, and problems with fonts. These, however, can all be fixed with the right software and expertise (the chapter on composition and graphics discusses using **color** [6.9], **image resolution** [6.11], and **fonts** [6.7] in more detail).

5. Binding

Finally, digital printing requires different binding equipment than does offset. There is no market where this is more obvious than the book market. Offset book printing is done on large presses that produce signatures of eight, sixteen, or thirty-two pages. These have to be folded, gathered (collated), and then bound. Digital book printing is done one, two, or three pages at a time, and they are delivered in sequence from the machine, so no collation is required. Traditional binding equipment handles many hundreds of books per hour. It is unlikely that the binding process for a digital printer, even a very fast roll-fed one, would ever need to exceed one hundred books per hour. And for a sheetfed machine, the number would be a fraction of that. The traditional bookbinding equipment is simply inappropriate.

Suppliers of binding equipment are belatedly responding to this need, especially in the key perfect-bound (softcover) area. There are expensive machines (several hundred thousand dollars) designed to handle the output of the roll-fed printers and the fastest sheetfed ones, and inexpensive ($15,000 and less) manual machines designed to handle book-on-demand binding from slower sheetfed devices. For the moment, though, there are no machines in the great gap between.

8.3 USES OF DIGITAL PRINTING

(I) DIGITAL PRINTING VS. OFFSET PRINTING

Most books and journals are printed via the offset process. (Although other traditional printing processes, including letterpress, gravure, and flexography can also be used, this happens only rarely. See the chapter on **composition and graphics** [6.13] for a discussion of these processes.) For some book printing, digital printing (which includes toner-based processes, ink-jet printing, and other processes that do not involve making a plate or other master) is starting to be an important competitor to offset. Important suppliers of toner-based machines for book printing include Xerox, Heidelberg, HP/Indigo, Oce, Xeikon, and IBM. At present there is only one significant supplier of ink-jet printing equipment suitable for books, Scitex Digital Printing, but others are preparing to enter this field.

To understand the developing competition between offset and digital printing, it is useful to consider the main characteristics that distinguish them today: quality, speed, flexibility, cost, and operational factors.

1. Quality

Though both toner and inkjet printing have historically been unable to match offset quality, that has recently changed. Now, the best of the toner-based "digital presses" can match the quality of most commercial offset work. Some low-speed inkjet devices are also capable of offset-quality printing, but for now they are far too slow to be used for book and journal printing and are used mostly for proofing.

2. Speed

Toner-based machines cannot, in general, match the speed of an offset press running continuously. But an offset press doing short runs spends more time during job-to-job changeover than it does actually printing, and in that situation a toner-based device can be as fast or faster. The high-speed inkjet printers from Scitex Digital Printing are much faster than the toner-based machines, and they can compete directly on speed (although not yet on quality) with offset presses.

3. Flexibility

Each approach offers specific kinds of flexibility. Offset printing currently offers more flexibility in terms of papers, inks (including spot colors and varnish), and binding options. Digital printing, of course, offers the flexibility of completely changing each page from one impression to the next.

4. Cost

Offset has a higher cost for starting up a job, because plates must be prepared and some paper is wasted during the initial run-up to the first good sheet. The binding process that goes with offset (print in signatures, fold, gather, bind, and trim) can also be more expensive than the binding process for digital printing, where large signatures and gathering are not required. But, at present, the cost per page of the actual printing is lower for offset, and this is the dominant cost in long runs. So short runs are more economical with digital printing, and longer ones are more economical with offset. The crossover point between the two methods is somewhere between 200 copies and 5,000 copies, depending on the exact equipment being compared.

5. Operational factors

Generally speaking, offset printing requires a highly-trained operator and a production (factory-like) working environment. The quality obtained from an offset press often depends on the skill of the operator. Digital printing tends to place less demands on the operator and to produce more uniform results. It can usually be done in an office-like environment. One implication of these differences is that highly-efficient, "lean" manufacturing of print is more likely to be accomplished with digital devices, given devices of comparable speed.

(ii) PRINTING ON DEMAND

The phrase "printing on demand" (***POD***) is used in various ways, but we will choose the most restrictive meaning: the printing of a document in response to the requirement of a specific end user. This definition makes clear the difference between POD and "short runs" (see separate discussion in the section on **short runs** [8.3.3]). Examples of POD would include the production of a manual for a bulldozer at the time when the bulldozer itself is produced, and the in-store printing of a book for a waiting customer. Printing on demand is one of two types of printing (the other is variable-data printing) that digital printing can address but conventional printing, which is restricted to producing multiple copies of the same document, cannot.

Most printing that is done in the office and the home could be classified as printing on demand. Items are printed as the need arises, and only the quantity that can be immediately used is produced—often just one copy.

But successful examples of printing on demand as a business are hard to find. The reason for this is not technical. The technology for printing on demand is exactly the same as it is for short runs: if you can print a run of fifity copies, you can print a run of one. Rather, the non-printing costs surrounding printing on demand (selling, processing the order, delivering the document, and billing) tend to make the whole process uneconomic. A particular printer may be able to produce a single copy of a paperback book at a cost of five dollars, for example, but if it costs the printer an additional five dollars to process the order and do the billing, the prospects of a profitable POD business are slim. And it is hard to fit a sales force paid on commission, which is often the case in the printing industry, into a print-on-demand model.

1. Books on demand: making it work

Because of the business issues discussed above, successful ***POD*** operations generally need to be highly automated and paid for in an efficient transaction.

On-demand production of individual books in the bookstore, if it ever becomes a commercial reality, would be an example. Here, the transaction would be handled by a clerk who is already handling many other sales, so the transaction costs would be low. As of late 2002, there is only one actual case of a book production device being installed in a bookstore—the machine is made by the Instabook Corporation, and the bookstore is Book Express in Cambridge, Ontario, Canada). Several companies are developing similar devices, and they should appear on the market over the next few years.

The engineering challenge of creating a self-contained book-production machine is formidable, but a greater obstacle may turn out to be setting up the infrastructure that allows such a machine to have access to all the books it might be called upon to print, while still guaranteeing the publisher a suitable commission and no loss of control over the intellectual property. This problem, like the analogous problem for e-books, is far from being solved. The case of Book Express is instructive: the Instabook machine seems to be working fine, but there very few books are produced on it because of the limited number of titles available.

In a few cases, print-on-demand book production has reached relatively large scale. The best example is Lightning Source, a subsidiary of the book distributor Ingram. Lightning Source has developed a very efficient system for producing books digitally. The work produced is a combination of true on-demand printing (one or two copies) and short runs. By mid-2002, the company had produced over three million books and had over 100,000 titles in its system.

(iii) SHORT-RUN PRINTING

Once the print run is underway, digital printing costs more per page than offset printing. The cost difference is very high (but declining every year) for color pages. It is small but still significant for black-and-white pages. This economic fact means that digital printing can be justified only in selected niches where other factors make up for the high cost per page.

Some costs are incurred in offset printing before the first good sheet comes off the press. Two such cost areas, the cost associated with platemaking and the cost associated with the wasted sheets used in getting the press "up to color," are avoided entirely in digital printing. A digital device needs no plate and switches instantly from one job to the next. Normally, neither paper nor time is wasted during the switch. This means that if the print run is short, the overall cost for printing a job digitally can be less than offset even though the cost per page is higher with digital printing.

For jobs involving fewer than 200 copies of a document, digital printing is almost always less expensive than offset. Depending on exactly what kinds of presses are being compared, digital printing can sometimes be competitive up to 1,000 copies or more. And sometimes factors apart from the printing itself (for example, the avoidance of signature gathering when printing documents with high page counts) can help justify digital printing, making it economical at still higher quantities.

1. Applications for short-run printing

Short-run digital printing has found application in a number of book-publishing niches, including:

1. COURSE PACKS AND CUSTOM TEXTBOOKS

These are materials selected by a professor for a specific course, and produced in the quantity required for the enrollment in that course.

2. ADVANCE COPIES AND REVIEW COPIES

These are copies of a book produced ahead of the primary (offset) print run and used for promotion. Publishers want their books to be reviewed to coincide with their availability in bookstores, and they want the sales force to have advance copies to obtain bookstore orders. In the past, these needs for early copies have been satisfied, if at all, with "bound galleys" produced on high-speed copiers. Digital printing provides a better solution.

3. BACKLIST BOOKS

If a book has slow but steady sales of a few dozen to a few hundred copies per year, the publisher will be reluctant to tie up capital in an offset print run that will take years to sell out. But short runs of twenty to fifty books make sense in this situation, even though the cost per book is comparatively high.

4. SMALL PUBLISHERS AND SELF-PUBLISHERS

For many small book publishers, and most self-publishers, a very short initial print run, using digital printing, is a wise choice. If the book sells in sufficient quantity, a switch to longer offset print runs can be made.

2. The success of short-run digital printing

In the twenty-year history of digital printing, by far the majority of the work produced has been in the short run category. The reasons for this are organizational rather than technical: short-run printing doesn't require new approaches to purchasing, sales, and job management (as POD does), and it doesn't require new skill sets and new ways of relating to the customer (as variable-data printing does). For printers (and their customers) who move from offset printing into short-run digital printing, the relationship and procedures can remain essentially unchanged.

With each passing year, new digital printing equipment is introduced that features higher speed and lower cost per page. This causes digital printing to become ever more competitive with offset. The longest economical short run will get longer and longer. Eventually (probably within a decade or two) there will be no difference in per-page costs, and digital printing will take over most of the work that is currently produced using offset presses. At that point, there will no longer be a reason to distinguish short runs from other press runs.

(iv) PERSONALIZED AND CUSTOM DOCUMENTS

Custom and personalized documents are an area of special strength for digital printing. Custom documents are those created for a small group of readers; personalized documents are each unique. It is practical to produce custom documents via offset printing for an audience of at least a few hundred, but smaller print runs call for digital printing. And of course, personalized documents (with a run length of one) require digital printing.

The actual printing of custom documents tends to be less important than the creation of the document content, which is where most of the costs tend to lie. There have been relatively few examples of successful business opportunities based on the production of custom books. The most notable one is custom college textbooks. Several publishers, beginning with McGraw-Hill's pioneering Primis project, established repositories of content in various fields of study. Professors get to select the content they want for their classes and have it produced as a custom book. The technology to do this has been around for almost two decades, and there are several successful examples, but this type of custom publishing is still not wide-spread. It turns out that establishing the repository (and especially, obtaining permissions and setting up royalty arrangements for all the content items) is the most difficult aspect. Another fairly common application of custom publishing is in creating spin-off products from a database. For example, the technique has been used to create industry-specific subsets of a catalog for a company to hand out at a specialized trade show.

Personalized printing is occasionally used for books. Sometimes, each copy of a custom-published textbook is imprinted with the student's name, which is a trivial example of personalization. Personalized children's books, with the use of the child's name and other information (e.g. the name of a pet, or the child's street), are common. Sometimes even the child's photo is scanned and used in illustrations throughout the book.

(v) THE ROLE OF THE WEB

Much has been made of the Web as a potential competitor to print publishing. But the Web is a key enabler for certain kinds of printing. As is noted in the discussion of **printing on demand** [8.3.2], the transaction costs for books printed on demand are a key factor in making the process profitable. Books ordered over the Web can have a very low transaction cost, since the customer does all the work of specifying the book, the delivery address, the payment option, and so on. Similarly, custom and personalized book production operations are much more likely to be profitable if the books are ordered via the Web.

The Web can also be helpful in streamlining the process leading up to the publication of a book. A notable example of this is the large number of book production services that have sprung up to serve the small publisher and self-publisher. These companies tend to use the Web for almost all of their interactions with customers, including bidding and contracts as well as the receipt of the finished manuscript and illustrations.

8.4 AVAILABLE PRINTING SYSTEMS

(I) SHEETFED PRINTERS

The most common digital printing devices print on individual sheets of paper. They aren't as fast as roll-fed devices (which are covered **below** [8.4.2]), but they are more flexible (since it is easy to switch paper stocks and sizes) and usually less expensive. The first digital printing devices to be commercialized were monochrome laser printers, in the late 1970s. Not until the 1990s did digital printing in color really become practical.

1. Monochrome sheetfed printing

In 1976, the first of two high-speed digital printing machines was introduced. The two were notably similar in their basic laser-printing functionality, but very different in their implementation. What's more, the segments of the digital printing market that were pioneered by each machine would remain separate for decades. Only now, two-and-a-half decades later, are they finally merging to forn a truly mass market.

1. IBM 3800

First to arrive was the IBM 3800. Its speed was astonishing for the time: 215 ppm. However, the 3800 made little impact on the publishing industry, for a variety of reasons—two of the major ones being insufficient resolution at 240 dpi and woefully inadequate fonts. The 3800, though, was not intended for publishing in the first place. It was intended to replace line printers attached to mainframe computers, a niche it filled quite well. At the time it appeared, all printed output in the mainframe environment was generated on mechanical impact printers—chain printers, train printers, flying-drum printers, and others of that ilk. They were noisy, inflexible, and often unreliable. During the machine's thirteen-year lifetime, IBM sold over 8,000 3800s. These trailblazing machines revolutionized the data center, and their offspring (right down to the HP LaserJet) revolutionized the office.

2. XEROX 9700

The other pioneering laser printer to appear in the 70s was the Xerox 9700, introduced in 1977. Unlike the IBM printer, the 9700 was not originally intended as a replacement for any particular machine. It printed 120 ppm at 300 dpi, and it could use any fonts that could be loaded into its memory (from 8KB to 32KB of bitmap memory, enough for 4 to 12 fonts at 10 point). One drawback was that a separate font was needed for each different size, and larger sizes took up more memory. Fonts could be downloaded along with the job to be printed.

The 9700 and its successors (down to the Apple LaserWriter) revolutionized document printing and publishing. Just as the IBM 3800 gave rise to an entire industry of laser printing for mainframe (and eventually office) output, the Xerox 9700 initiated an industry of laser printing for documents and publishing. (To this day, in the high-speed printing market, the convergence of the two segments is still not quite complete.)

Like all machines of its time (and most others for a decade to come), the 9700 was basically a text-only machine. The only graphics it could handle were small logos and crude line art assembled from special "fonts." These were limitations imposed by the technology of the day, in particular, the high cost of random-access memory (RAM). The 9700 was ideal for textual documents of high value and variable content, and it turned out that the insurance industry was its natural home. It could print out single copies of custom policies, replacing a pick-and-assemble manual process.

But the 9700 and the machines that followed it were too limited and required too much computer infrastructure to be used for ordinary publishing applications.

3. POSTSCRIPT AND DOCUTECH

Until the arrival of ***PostScript*** laser printers on the market in 1985, each brand of laser printer had its own interface language or "page description language," and software to drive laser printers was complex and expensive. PostScript would eventually put an end to the proliferation of proprietary languages, but it took years before computer performance increased to the point that driving a high-speed printer with PostScript was practical.

In 1990, Xerox announced the ***DocuTech***, the machine that dominated publishing applications of monochrome digital printing throughout the 1990s. This 120 ppm digital printer was initially marketed as a very sophisticated copier, but Xerox soon introduced a PostScript capability for it and it was increasingly used as a high-speed digital printer rather than a copier. With the DocuTech, Xerox began to displace large amounts of monochrome offset printing. It introduced new models in various speed ranges, culminating with the DocuTech 6180 at 180 ppm. It is still the fastest sheetfed digital printer capable of duplex (two-sided) printing.

The Xerox DocuTech didn't have much competition, though Kodak (first on its own and later in partnership with Heidelberg) did offer a 110 ppm device. The higher-speed DocuTech models were without competition until the 150 ppm Heidelberg Digimaster 9150 and the 155 ppm Oce VarioPrint 5160 were finally introduced in the fall of 2002.

4. THE MARKET TODAY

Xerox continues to dominate the high-speed monochrome sheetfed market. Apart from Heidelberg (whose machines are also sold by Canon, IBM, Ikon, and Danka), there are a number of vendors (led by Canon and Ricoh) with laser printers at 100 ppm and below, all looking for an opportunity to take some of the remaining monochrome business from offset printing (or from Xerox and Heidelberg). Some of these machines are very aggressively priced, and it is clear that both the initial investment and the cost per page of high-speed digital printing will decline sharply.

The DocuTech and the machines that compete with it owe a lot to copier technology. They all use photosensitive drums or belts, and they use either a laser or LEDs to create the image to be printed. Most of the manufacturers (Heidelberg being the major exception) are very active in the copier market as well.

There is one monochrome sheetfed machine worthy of mention that is outside the mainstream markets: the Delphax Imaggia. This is a specialized machine for check printing. It is actually faster than any other sheetfed printer, but it prints only on one side. It is capable of printing with magnetic ink (a requirement for checks) and it can handle an extremely wide range of paper weights. It uses Delphax' unique imaging technology, called electron-beam imaging.

2. Color sheetfed printing

The earliest color digital printers were basically color copiers with a computer interface. Their low speeds and modest quality, combined with a relatively high cost per page, meant that they could not be used for any "publishing" applications except early-stage proofing. In 1993, two companies—Indigo and Xeikon—began to change this. Xeikon's machine was roll-fed and is discussed **below** [8.4.2]. Indigo's E-Print 1000 was a 17 ppm printer using liquid toner (the company called it "Electro-ink"). From the start, it produced quality that could compete for many offset jobs.

During the course of the 1990s, copier companies gradually expanded the market for office color printing. Canon in particular found a corporate market for its CLC series of copier-based machines which are fast (up to 50 ppm) but which do not produce images whose quality can match Indigo's. The Canon machines are widely used in corporate and quick-print environments.

Xerox, too, offered copier-based printing to the corporate market. But with its introduction of the DocuColor series in the late 1990s, it began to focus more seriously on high-quality publishing markets in which it would be competing with Indigo. It took a further step in this direction with its iGen3 printer, a machine designed specifically to compete with offset printing. It was shown as a prototype at Drupa, the major international printing trade show, in 2000, and officially announced the following year. As of the end of 2002, volume production of this 100 ppm machine is just getting started.

In parallel with Xerox' development of the iGen3, Heidelberg and Kodak were also working on a machine to address offset markets, the NexPress 2100. It was first shown at Drupa 2000, and shipments of this 70 ppm machine began in 2001.

Meanwhile, Indigo had not been sitting still. It had made speed and quality improvements and had introduced a roll-fed model and a machine designed for printing on flexible plastic. But the biggest milestone for Indigo came in the fall of 2001, when it was purchased by Hewlett-Packard. This provided new financial and marketing resources, and it held the promise of streamlined, low-cost manufacturing in the future.

Thus, by the fall of 2001, the basic structure of the sheetfed digital printing market of today was in place. HP Indigo, Heidelberg, and Xerox are competing with each other for a share of the market that has traditionally been held by offset printing. These three vendors have deep pockets, broad market access, and expensive machines with very high image quality. Then comes a second group of companies, dominated by the copier vendors Canon, Sharp, Konica, Oce, and Minolta, among others. They are focused on the corporate market and the lower-quality segments of the offset market.

These are all toner-based machines. (So far, no sheetfed inkjet machines are fast enough to compete for production work, although they are widely used for low-volume tasks such as proofing and book-cover printing.) Almost all of them use a photosensitive drum or belt, with a laser or LED array to create the image. The exception to this is the Oce CPS700, which uses circular electrodes embedded in its imaging drum to attract toner. This machine, which is also unusual in utilizing seven different colors of toner, produces excellent images. But its market has been limited so far by its relatively low speed (25 ppm) and high price.

(ii) ROLL-FED DIGITAL PRINTING

For really fast printing, a roll-fed (or "web," to use the traditional printing term) machine is needed. Paper-handling difficulties put an upper limit on how fast a sheetfed printer can be. So far, no sheetfed machine has been able to print faster than 180 letter-size pages per minute. But roll-fed machines can go much faster. Some of them print at speeds of 750 or 1,000 feet per minute which (depending on the maximum width of the roll) can be the equivalent of over 2,000 letter-size pages per minute.

But the use of paper in rolls has its disadvantages as well. Only one kind of paper can be loaded at a time (sheetfed machines can have as many kinds as there are paper drawers), and changing paper generally requires heavy equipment and several minutes of downtime. Once the new roll is loaded, though, the machine will be able to run for hours without attention.

Roll-fed machines, especially the fastest ones, present special data-processing challenges. These days, most publishing-oriented pages are printed from ***PostScript*** or ***PDF*** files. But it can take a second or two for a fast processor to take each PostScript or PDF page and "rasterize" it—break it down into the individual dots that the printer will ultimately image. Photographs, in particular, require a lot of processing power. If the processing takes, say, a second per page, but the printer is capable of printing ten or more pages per second, then the rasterizing process becomes a major bottleneck. The speed of transfer of data from the host computer to the imaging part of the printer can also be a bottleneck.

Various solutions have been devised to help with this—multiple processors working in parallel, parallelism in the transfer of data, rasterization and compression of images ahead of time, and so on. The use of simpler languages than PostScript and PDF can also help. Xerox Metacode, IBM AFP/IPDS, and Hewlett-Packard PCL are all much faster to rasterize than PostScript and PDF (although they are far less flexible), and so they are still widely used, especially in non-publishing applications (e.g.; statement printing). This problem will gradually solve itself, since the processing speed of the average computer chip is increasing faster than the printing speed of the fastest printers. Eventually, the chips will be fast enough to keep up without difficulty.

While sheetfed printers have a lot in common with copiers, and are frequently based on copier designs, roll-fed machines are much more diverse. They use a variety of printing technologies that are not found among the sheetfed machines.

1. Monochrome machines

Most roll-fed printers installed so far are monochrome machines. They have been used for many years in statement printing (e.g., phone bills, credit card statements, brokerage account reports). But centralized statement printing is not a growth market, and all the companies that supply roll-fed machines have been seeking out additional markets. In recent years, a lot of emphasis has been placed on short-run printing of books. Other important markets have included personalized direct mail and printing of numbered or personalized documents (such as tickets and credit cards).

1. IBM

For many years, IBM has been active in this market. At one time, it was a dominant force in data-center printing. In recent years, IBM's Printing Systems division has focused increasingly on sales where printing is only one part of a solution that also involves consulting, software development, and integration. This has often meant that IBM has chosen not to bid on projects where the focus was the straightforward purchase of a printer. IBM's latest roll-fed system, the 4100, prints at up to 762 ppm and is a good fit for book printing.

2. OCE

IBM's narrow focus has opened up some opportunities for Oce (formerly Siemens), which originally made its mark selling IBM-compatible printers in the mainframe environment. It now sells roll-fed machines into a variety of applications, including book printing, and has become the leading vendor of roll-fed printers. Oce's VarioStream products run at up to 1,273 ppm and are used in a number of book-on-demand production facilities.

3. XEROX

Xerox, though it concentrates on sheetfed machines, also offers roll-fed ones. It offered machines built by Delphax (see below [8.4.2.1.5]), of which it was part owner, then switched to machines built by a Japanese subcontractor when Delphax was bought by Check Technology. The Xerox CFD models print at up to 1000 ppm.

IBM, Oce, and Xerox all offer roll-fed machines that use the traditional approach of creating the image on a light-sensitive drum or belt with a laser or LED array. But other roll-fed machines use more unusual technologies.

4. NIPSON AND MAGNETOGRAPHIC PRINTING

Nipson, a monochrome specialist, uses a unique technology called "magnetography" in its printers. A writing head, analogous to the head of a disk drive, creates magnetic spots on the surface of steel drum. Toner containing iron particles sticks to the magnetized spots, forming the image. The toner is then transferred to paper. Nipson's fastest machine, the Varypress, prints at 1,616 ppm. Nipson's main specialty is printing on all sorts of substrates (such as plastic credit cards and heat-sensitive materials) that other printers have difficulty with. But it is also active in direct mail printing and, in recent years, has begun paying more attention to book printing, especially text-only books. Nipson's magnetographic technology is unsuitable for color printing, since the materials that would be needed to make toners that have both the necessary magnetic characteristics and the required colors do not exist.

5. DELPHAX AND ELECTRON-BEAM IMAGING

Delphax was originally a joint venture of Xerox and two other companies, and then became a Xerox subsidiary. It was sold to Check Technologies in December 2001. Its printers utilize a unique electron-beam imaging technology. An alternating current knocks electrons loose from the air in tiny holes in the print head. The electrons are deposited on the surface of the nearby imaging drum or belt, and the resulting negative charge attracts toner. The technology is fast, simple, and rugged. Delphax has incorporated it into a sheetfed check printer (the Imaggia, mentioned **earlier** [8.4.1.1.4]) and a line of roll-fed devices, the fastest of which prints at 1300 ppm. The machine would be suitable for book printing. All the Delphax models are monochrome, but the company has begun development on a full-color machine for introduction in 2005.

6. SCITEX DIGITAL PRINTING AND VERSAMARK

The fastest of the monochrome printers is the Versamark press from Scitex Digital Printing. This is a high-speed inkjet machine (up to 750 feet per minute, or over 2,000 pages per minute). Its key features are its speed and low cost per page. Its print quality is limited, but its other features make it well suited to the production of materials where the highest quality is not an issue. And a new generation of printhead technology, shown in prototype form in the fall of 2002, will close part of the gap in image quality between this technology and the toner-based machines.

7. ELCORSY ELCO 400

A newcomer, just reaching the first customer site as this is written, is the Elcorsy Elco 400. This machine also features high speed (400 feet per second, or about 870 letter-size pages per minute) and low cost per page. The novel technology used in this machine is called "electro-coagulation." A special fluid, containing a pigment and a polymer, is passed between a rotating metal drum and an array of electrodes that is very close to the drum's surface. Whenever one of the electrodes is turned on, a spot of pigment is deposited on the drum. As the drum turns, it comes into contact with paper and the pigment is transferred. The technique should be capable of quite good quality. Early samples showed streaking problems, but these seem to have been fixed by a redesign of the imaging head. A special attribute of this technology is that it is able to print ***grayscale*** (8-bit) images, thus achieving high image quality compared to other digital printers. The machine's initial installations will be monochrome (or monochrome with spot color), although the technology is suitable for color printing as well.

2. Color machines

Two of the vendors of monochrome roll-fed systems (Scitex Digital Printing and Elcorsy) also have (or, in Elcorsy's case, plan to have) color machines. The Scitex Business Color Press is the speed champion, at 500 feet per minute. Elcorsy's competing product, when available, will run at 400 feet per minute. Both are far faster than the toner-based competition. The image quality of the Scitex product is significantly inferior to most toner-based printing, but may still be suitable for some books. The Elcorsy device is expected to have good image quality, and the prototype machine is promising, but the image quality issue will only be completely resolved by production use at customer sites.

Meanwhile, toner-based roll-fed color machines are available from Xeikon and HP Indigo. Xeikon started out in 1993 with a roll-fed color printer and, after dabbling briefly with a sheetfed model, has decided to remain a roll-fed specialist. Its two main products are the DCP 500D and the DCP 320D, which accept maximum paper widths of 20 inches and 12.6 inches, respectively. They each use eight imaging stations (four on each side of the paper) to print both sides simultaneously at up to 130 ppm. In digital book-printing operations, the Xeikon printers (which are also sold under the IBM and MAN Roland names) are often used for printing color covers and jackets.

HP Indigo initially offered only sheetfed printers, but has added the roll-fed w3200 model that prints 133 ppm. Among the roll-fed printers, it is clearly the leader in image quality.

For now, these are the only vendors active in color roll-fed printing for publishing applications. But others will probably enter the market. Several inkjet specialists have shown interest, and one of them, Aprion, showed a mockup of a self-contained book-on-demand printing system at a trade show in 2000. Another inkjet vendor, Dotrix, is actually selling a fast roll-fed machine that could in principle be adapted for publishing use. But so far, the company is marketing it as a machine for printing labels, plastic cards, packaging, wall coverings, and other non-publishing applications.

(iii) CLUSTER APPROACHES

It is possible to use several relatively slow printers, working simultaneously, to duplicate the throughput of a single fast machine. There are several reasons why this approach is attractive. You can start small (with one or two printers) and build capacity incrementally as required. You have a measure of redundancy, so the failure of a single printer does not stop the whole production operation. You may be able to save money, both on your initial investment and on your running costs, because of what might be called a "dis-economy of scale" in the printer market. For example, the price of a 120-ppm printer is more than the cost of four 30-ppm printers, and the cost per page is often higher on the faster printer as well.

One vendor in particular, T/R Systems, has specialized in cluster systems and has sold many of them. Some have been used in book printing. But interest in cluster systems has declined in recent years. One reason is that both the initial investment and the cost per page of fast printers has been declining, making the economics of clustering less attractive. But a more important reason is the new class of enterprise printer-management software that has become available. This software monitors printer use and printer failures throughout an enterprise, routing jobs to an alternative printer in case of machine failure, splitting large jobs among multiple machines, and handling internal accounting and billing. It provides most of the features of cluster printing as a subset, so there are fewer potential customers for software that is restricted to cluster printing. In addition to these enterprise-wide packages, many of the printer vendors now offer print-management software of their own that includes clustering features.

8.5 SALES CHANNELS FOR DIGITAL BOOK PRINTING

As digital printing becomes a common approach to book production, it is changing the business practices of the book-printing industry. Existing book printers find they have to make adjustments to allow for more frequent, but shorter, print runs. And many companies that were not previously involved in book printing are joining the competition. The new digital technologies have changed the market at two levels: they have led to a dramatic drop in the cost of the equipment required to become a book printer (it can now be done for a cost of under $20,000 if you are producing only a few dozen copies a day), and the minimum length of a print run has also dropped (in some cases, to a single copy).

(i) BOOK AND JOURNAL MANUFACTURERS

Many existing book and journal printers are embracing the digital technology. Typically, they do not offer a true "book-on-demand" service. Rather, they focus on short-run printing of from fifty to five hundred copies. Short print runs can often be handled within their existing production and job-tracking processes, and the adjustments to their sales and billing processes can be kept to a minimum. They can sell digital printing as an additional service to their existing book-publisher customers (who can still be expected to buy offset printing when the runs are longer).

(ii) BOOK WHOLESALERS AND DISTRIBUTORS

In some cases, book distributors and wholesalers are in a good position to implement digital book-on-demand services. They have to keep books in stock to supply their bookstore customers, so they know when a book goes out of print. They also know when there is demand for a book that is no longer available. So they are in a good position to sell publishers a service that would allow small numbers (or single copies) of slow-moving books to be produced in response to market demands. The most aggressive efforts in this direction to date have been Ingram's Lightning Source operation (based on IBM equipment) in the U.S. and Libri's similar service (based on Xerox equipment) in Germany. Both are set up to print single copies efficiently, though they also offer short runs.

(iii) VANITY PRESSES AND SELF-PUBLISHING SERVICES

Vanity presses (publishers who are paid by authors to produce their books) have been around for centuries. But the traditional book-printing technology has not been a good match for the very short runs that self-published authors often require. Vanity presses have therefore embraced digital technology, and today digital printing is often used for the books they produce. In addition, dozens of book-production services for self-publishers have sprung up in recent years, based on the availability of digital printing. These services—which are often indistinguishable from vanity presses, except that they generally do