Inv.-Nr.: 13944

Content Management Bible, 2nd Edition

Bob Boiko

WILEY
Wiley Publishing, Inc.

Contents

Foreword ix

Preface xi

Acknowledgments xvii

Introduction xxxix

Part I: What Is Content? 1

Chapter 1: Defining Data, Information, and Content 3

What Is Data? 4
Content Is Not Data 5
Content Is Information Put to Use 7
Content Is Information Plus Data 11
From Data to Content and Back 11
Summary 12

Chapter 2: Content Has Format 13

Storage Formats: Storing Information 13
Rendering Format: Presenting Information 14
Dealing with Formatting 15
Categorizing Formatting 17
- Formatting for effect 17
- Formatting by method 18
- Formatting by scope 18

Summary 19

Chapter 3: Content Has Structure 21

Structure Is Important 21
Structure Can Be Difficult to Create 24
How to Categorize Structure 26
- Structure by purpose 27
- Structure by type 27
- Structure by scope 28

Summary 29

Chapter 4: Functionality Is Content, Too! . . . 31

What Is Functionality? . . . 31
Monolithic versus Mix-and-Match Functionality . . . 32
Functionality Interspersed with Information . . . 33
Managing Functionality Like Information . . . 35
Publishing functionality on the Web . . . 36
Where information and functionality diverge . . . 38
Summary . . . 40

Chapter 5: But What Is Content Really? . . . 41

Content, Context, and Meaning . . . 41
Creating Context Rules . . . 43
Content Organization Starts with Purpose . . . 45
Content Is Named Information . . . 47
From Data to Wisdom . . . 49
Data is raw and discrete . . . 49
Data is nondiscursive and out of context . . . 51
Information is processed and continuous . . . 52
Information is discursive and full of context . . . 54
Knowledge and wisdom can be information . . . 57
Why Does Text Get All of the Attention? . . . 59
Text is codified language . . . 60
Text has the lead . . . 61
Summary . . . 62

Part II: What Is Content Management? 63

Chapter 6: Understanding Content Management . . . 65

Defining Content Management . . . 65
CM Is Distributing Business Value . . . 67
CM Is a Balance of Organizational Forces . . . 69
CM Is the Combination of Content-Related Disciplines . . . 70
CM Is Collection, Management, and Publishing . . . 72
CM Is a Computer Infrastructure . . . 74
The static Web site . . . 75
The dynamic Web site . . . 75
The Web CMS . . . 77
The full CMS . . . 79
The enterprise CMS . . . 81
The Content Management Industry . . . 82
Summary . . . 83

Chapter 7: Introducing the Major Parts of a CMS . . . 85

A CMS Overview . . . 86
The Collection System . . . 87
Authoring . . . 88
Acquiring . . . 89

Converting . . . 92
Aggregating . . . 93
Collection services . . . 99
The Management System . . . 100
The repository . . . 100
The administration system . . . 103
The workflow system . . . 105
Connections . . . 105
The Publishing System . . . 106
Publishing templates . . . 106
Publishing services . . . 108
Connections . . . 108
Web publications . . . 109
Other publications . . . 111
Summary . . . 111

Chapter 8: Knowing When You Need a CMS . . . 113

Gauging the Amount of Content . . . 114
Managing the Size of the Contribution Base . . . 115
Anticipating the Amount of Change . . . 117
Knowing the Number of Publications . . . 119
Estimating Complexity . . . 121
Vacation company "A" . . . 123
Vacation company "B" . . . 124
Vacation company "C" . . . 125
Vacation company "D" . . . 126
Evaluating your own need . . . 128
Summary . . . 129

Chapter 9: Component Management versus Composition Management . . . 131

CM Systems Can Be Modular or Linear . . . 131
Component CM Systems . . . 133
Composition Systems . . . 135
Collection in a composition system . . . 136
Management in a composition system . . . 137
Publication in a composition system . . . 138
Schema-Driven Systems . . . 141
Collection in a schema-driven system . . . 143
Management in a schema-driven system . . . 144
Publishing in a schema-driven system . . . 144
Which System Is Right for You? . . . 145
Summary . . . 146

Chapter 10: The Roots of Content Management . . . 147

The Knowledge Base of the Publishing Industry . . . 148
The publication . . . 148
Content collection . . . 149
Abstracting the process . . . 149

The Principles of Document Management . . . 151
Files contain content . . . 152
Files store BLOBs . . . 153
Publications are files . . . 153
Document management systems versus content management systems . . . 154
The Limits of Information Technology Groups . . . 155
IT departments have avoided the Web . . . 156
Content begins where IT groups leave off . . . 156
The Legacy of the Multimedia Industry . . . 157
Electronic publications . . . 159
End-to-end content management . . . 159
The Challenges of Technical Communication . . . 160
Huge information bases . . . 160
Simultaneous publications . . . 162
Communication Theory . . . 163
The Traditions of Library and Information Science . . . 164
Information behavior . . . 164
User services . . . 165
Knowledge representation . . . 165
Information retrieval . . . 166
The Technology of Software Development . . . 168
Collection technologies . . . 168
Management technologies . . . 168
Publishing technologies . . . 168
The functionality in electronic publications . . . 169
The Premises of Marketing . . . 171
Summary . . . 172

Chapter 11: The Branches of Content Management . . . 173

Personalization . . . 174
What is personalization? . . . 174
Content management underlies personalization . . . 176
Advanced Web Sites . . . 177
Multiple Publications . . . 179
E-commerce . . . 182
Catalogs and content management . . . 183
E-commerce functionality and content management . . . 184
Knowledge Management . . . 185
What is knowledge management? . . . 185
Knowledge is content to manage . . . 186
Online Communities . . . 187
What is a community? . . . 188
How are online communities constructed? . . . 188
Other Kinds of Management . . . 193
Digital asset management (DAM) . . . 195
Learning object management (LOM) . . . 195
Source management . . . 195
Digital records management . . . 196
Digital rights management (DRM) . . . 197
Summary . . . 198

Part III: Doing Content Management Projects 199

Chapter 12: Doing CM Projects Simply 201

Why Create a Minimal CMS? 201
Staffing on a Shoestring 203
The business person 203
The content person 204
The publications person 204
The technology person 205
Getting Ready for the Project 205
Exploring the organization 206
Finding the right project 207
Finding the right sponsors 207
Getting a mandate 208
Key planning deliverables 209
Doing Design 209
Getting minimum requirements 209
Logical design essentials 210
Key design deliverables 210
Implementation 211
How low-tech can you go? 211
Slimming down product selection 214
Saying no 214
Key implementation processes 215
Key implementation deliverables 216
Deployment 217
Key deployment processes 217
Key deployment deliverables 218
Summary 218

Chapter 13: Staffing a CMS 219

About CMS Jobs 219
Managers 221
Content manager 221
Project manager 223
Production manager 223
Business Analysts 224
Information Architects 225
Content analyst 226
Metator 227
Infrastructure Staff 228
CMS administrator 229
Deployment analyst 230
Trainer and documentation specialist 231
Software Developers 231
Software analyst 231
Template and CMS developer 232
Custom application developer 232
Software integrator 233
Test analyst and test engineer 233

Technische Universität Darmstadt
Fachgebiet Sportwissenschaft

Publications Staff 234
Publication analyst 234
Publication designer 235
Page developer 236
User interface specialist 237
Content Processing Staff 237
Conversion analyst 237
Tool developer 238
Content processor 239
Content QA specialist 239
Content Creation Staff 239
Acquisitions manager 240
Traffic cop 240
Writers and other content creators 240
Editor 241
Summary 241

Chapter 14: Working within the Organization 243

Content Management and the Organization 243
Content Management versus Other Systems 247
Tracking Information Flow in the Organization 247
Understanding your information 249
Understanding your functionality 249
Understanding Organizational Roles 251
Business units generate value 251
Editorial teams unify content 251
Marketing teams direct and unify publications 252
IT groups build and maintain infrastructure 253
What Can Be Shared? 253
Product 253
Code 254
Content 254
Publications 255
How do you decide? 255
Exploring Organizational Models 256
Collection variations 257
Publishing variations 258
Management variations 261
Using Functional Collection and Publishing 263
Organizing collection systems and publications by type 263
Creating functional teams 264
Identifying Your CMS Hurdles 267
Summary 270

Chapter 15: Getting Ready for a CMS 271

Readiness Jumpstart 271
Understanding the CMS Project Process 272
Techniques for Getting the Job Done 274
Start with the project team 275
Look for pain in the organization 277
Assess the current mandate 278
Assess the organization's assumptions 278

Taking Stock of the Deliverables . . . 281
The readiness assessment . . . 281
Document inventory and analysis . . . 283
The state-of-the-content system report . . . 284
The education plan . . . 285
A preliminary project plan . . . 286
A risk assessment . . . 287
Taking Stock of Your Staffing Needs . . . 288
Summary . . . 288

Chapter 16: Securing a Project Mandate . . . 289

Mandate Jumpstart . . . 289
What to Agree on . . . 290
Techniques for Getting the Job Done . . . 291
Recognize sponsors . . . 291
Learn about your sponsors . . . 293
Taking Stock of the Deliverables . . . 294
Sponsor profiles . . . 294
A hierarchy of issues . . . 295
Notes and minutes . . . 295
The mandate process . . . 296
The project mandate . . . 299
Taking Stock of Your Staffing Needs . . . 302
Summary . . . 302

Chapter 17: Doing Requirements Gathering . . . 303

Requirements Jumpstart . . . 303
What Are Requirements? . . . 304
Techniques for Getting the Job Done . . . 304
The requirements process . . . 305
The requirements . . . 306
Approaching the organization . . . 310
Taking Stock of the Deliverables . . . 311
The requirements plan of attack . . . 312
The requirements document . . . 312
Taking Stock of Your Staffing Needs . . . 313
Summary . . . 313

Chapter 18: Doing Logical Design . . . 315

Logical Design Jumpstart . . . 315
What Is Logical Design? . . . 316
Techniques for Getting the Job Done . . . 318
Why do logical design? . . . 319
From business to system . . . 320
Iterating through your design . . . 322
Triangulating on constraints . . . 323
Managing the details . . . 325
Taking Stock of the Deliverables . . . 325
The design plan of attack . . . 326
The collection design document . . . 327
The management design document . . . 328
The publication design documents . . . 330

An audience analysis . . . 332
A localization plan . . . 333
A risk assessment plan . . . 335
A revised project plan . . . 336
An executive summary . . . 337
Taking Stock of Your Staffing Needs . . . 338
Summary . . . 339

Chapter 19: Selecting Hardware and Software . . . 341

System Selection Jumpstart . . . 341
The Product Paradox . . . 343
Build, Buy, or Rent? . . . 348
Building a CMS . . . 348
Buying a CMS . . . 349
Renting a CMS . . . 349
Techniques to Get the Job Done . . . 351
How to select decision makers . . . 351
How to select a product . . . 353
Taking Stock of the Deliverables . . . 364
Product files . . . 364
The selection criteria and RFP . . . 364
The score card . . . 365
The design diagrams . . . 365
The decision report . . . 366
Risk assessment update . . . 366
Project plan update . . . 367
Taking Stock of Your Staffing Needs . . . 367
Sorting Through the CMS Selection Criteria . . . 369
Business criteria . . . 370
Overall criteria . . . 371
Collection criteria . . . 377
Management criteria . . . 382
Publishing criteria . . . 390
Summary . . . 397

Chapter 20: Implementing the System . . . 399

Implementing the System Jumpstart . . . 399
Looking at the Process So Far . . . 401
Looking at the Project So Far . . . 401
Techniques for Getting the Job Done . . . 402
Cutting back . . . 402
Do you have one project or many projects? . . . 403
Detailing the implementation process . . . 405
Taking Stock of the Deliverables . . . 418
Collection specifications . . . 420
Management specifications . . . 421
Publication specifications . . . 422
The project plan . . . 424
The risk assessment plan . . . 426
Taking Stock of Your Staffing Needs . . . 427
Summary . . . 427

Chapter 21: Rolling Out the System 429

Rolling Out the System Jumpstart 429
What Is Deployment? 430
Techniques to Get the Job Done 431
Creating documentation 431
Doing training 437
Powering up the system 440
Preparing content 444
Revising the system 444
Taking Stock of the Deliverables 446
The deployment specification 446
The staffing plan 447
The training plan 447
The documentation plan 448
The maintenance plan 449
Summary 450

Part IV: Designing a CMS 451

Chapter 22: Designing a CMS Simply 453

Logical Design Essentials 453
The Entities at a Glance 454
Logical Design: An Example 455
PLAN International—an example organization 455
A very simple logical design 455
Getting Beyond the Simple Logical Design 458
Summary 458

Chapter 23: The Wheel of Content Management 459

The Content Management Entities 459
Introducing the CMS wheel 461
Entities add a layer of abstraction 463
Goals and requirements 464
Audiences 468
Publications 471
Content types 474
Authors 478
Acquisition sources 480
Access structures 483
Workflow and staffing 484
Think, Plan, Integrate 488
Summary 489

Chapter 24: Working with Metadata 491

What Is Metadata? 491
What does meta mean? 491
What does metadata mean? 492
What does metatorial mean? 494

The narrow view of metadata . . . 496
The wide view of metadata . . . 496
Metadata and content management . . . 497
Understanding the Types of Metadata . . . 498
Structure metadata . . . 498
Format metadata . . . 500
Access metadata . . . 502
Management metadata . . . 503
Inclusion metadata . . . 504
Categorizing Metadata Fields . . . 506
Metatorial Processing . . . 508
The metator . . . 510
The metatorial guide . . . 511
Localizing Metadata . . . 514
Summary . . . 515

Chapter 25: Cataloging Audiences . . . 517

Cataloging Audiences Jumpstart . . . 517
Serving versus Exploiting an Audience . . . 518
What Is an Audience? . . . 520
Audiences and communicators . . . 520
Audiences and marketing . . . 521
Audiences and users . . . 522
How many audiences do you have? . . . 522
Audiences and Localization . . . 523
What is localization? . . . 524
What are your localities? . . . 525
Audiences and locality . . . 526
What gets localized? . . . 526
Localization and content management . . . 527
An Example Audience Set . . . 528
Analyzing Audiences . . . 529
Think . . . 529
Plan . . . 530
Integrate . . . 535
Summary . . . 536

Chapter 26: Designing Publications . . . 537

Analyzing Publications Jumpstart . . . 537
What Is a Publication? . . . 538
Publication purpose . . . 539
Publishers . . . 541
Audiences . . . 542
Messages . . . 543
Authorship . . . 545
Publication format . . . 546
Publication structure . . . 546
Publications can be good or bad . . . 547
Publication gotchas . . . 549

Analyzing Publications . . . 552
Think . . . 552
Plan . . . 553
Integrate . . . 557
Summary . . . 558

Chapter 27: Designing Content Types . . . 559

Content Type Jumpstart . . . 559
The Idea of a Content Model . . . 560
What Are Components? . . . 563
Components are like objects . . . 564
The basic unit of content management . . . 565
How do you divide content? . . . 566
Components versus pages . . . 568
Content has types and components . . . 569
Content Types Have Elements . . . 570
Elements have types and values . . . 571
Unique identification . . . 572
Locality in components . . . 572
What Do Components Look Like? . . . 573
Components in flat files . . . 573
Components in structured files . . . 574
Components in relational databases . . . 575
Components in object databases . . . 575
The Relationship between Functionality and Components . . . 576
A functionality component . . . 577
Functionality content types and components . . . 580
Analyzing Content Types . . . 581
Think . . . 582
Plan . . . 582
Integrate . . . 589
A Sample Set of Content Types . . . 590
Affiliations . . . 592
Locations . . . 593
Solutions . . . 593
Offering . . . 594
Annual Reports . . . 595
SEC Filings . . . 596
Events . . . 596
Press Releases . . . 597
Newsletters . . . 598
Outside News . . . 599
Outside Links . . . 600
Articles . . . 601
FAQs . . . 601
Customers . . . 602
Jobs . . . 603
Downloads . . . 604
Bio . . . 604
Standard management elements . . . 605
Summary . . . 608

Chapter 28: Accounting for Authors 609

Authoring Jumpstart 609
What Is Authoring? 610
Who can become an author? 611
Harvesting creative product 611
Authoring tools 612
Authoring prior to the CMS 612
The Author's Attitude 613
Changing the author 613
Changing the content 615
Your Attitude Toward Authors 615
Technical savvy 616
Influence 618
Incentives 618
A savvy/influence matrix 619
Analyzing Authors 621
Think 621
Plan 621
Integrate 627
Summary 627

Chapter 29: Accounting for Acquisition Sources 629

Acquisition Jumpstart 629
What Is Acquisition? 631
Found sources 631
Syndicated content 632
To syndicate or connect? 635
Acquiring functionality 635
Is It an Author or an Acquisition Source? 637
Analyzing Sources 638
Think 638
Plan 639
Integrate 645
Summary 645

Chapter 30: Designing Content Access Structures 647

Access Structure Jumpstart 647
Understanding Access Structures 650
Publication navigation versus CMS access structure 652
The content domain 653
Hierarchies 656
Indexes 659
Cross-references 662
Sequences 666
Full-text search 668
Analyzing Access Structures 669
Think 669
Plan 670
Integrate 680
Summary 681

Chapter 31: Designing Templates 683

Templating Jumpstart 683
Publications and Templates 684
Bridging worlds 685
Mixing the static and the dynamic 687
Building pages 689
Creating a system of publications 689
Using a template processor 690
Understanding template logic 691
Using templates within templates 692
Web templating 695
Print templating 703
Fax templating 706
E-mail templating 707
Analyzing Templates 707
Think 707
Plan 708
Integrate 715
A Sample Set of Templates 715
Component templates 717
Navigation templates 723
Summary 731

Chapter 32: Designing Personalization 733

Personalization Jumpstart 733
Publications and Personalization 734
Personalization and the audience 734
Personalization and components 736
Personalization and rules 737
Personalizations in templates 738
Customization versus personalization 740
Dynamic and static personalization 742
Analyzing Personalization 742
Think 743
Plan 743
Integrate 752
Summary 753

Chapter 33: Designing Workflow and Staffing Models 755

Workflow Jumpstart 755
Understanding Workflow 757
Workflow triggers 761
Workflow objects 762
A perspective on workflow 764
Push-and-pull workflow 765
Introducing Tasks, Jobs, and Steps 766
Staff modeling 767
Staff, jobs, and tasks 768
Task-time calculations 770
Staffing calculations 771
Fudge factors 772

Fitting Localization into Your Workflow . . . 773
Analyzing Workflow . . . 777
Think . . . 777
Plan . . . 778
Integrate . . . 784
Analyzing Staffing . . . 784
Think . . . 785
Plan . . . 785
Integrate . . . 788
Summary . . . 788

Part V: Building a CMS 791

Chapter 34: Building a CMS Simply . . . 793

What Is Physical Design? . . . 793
Physical Design Essentials . . . 794
A Simple Physical Design Process . . . 796
Introducing the Technology Taxonomy . . . 798
An index to the technology taxonomy . . . 799
Summary . . . 803

Chapter 35: What Are Content Markup Languages? . . . 805

A Brief and Selective History of Markup Languages . . . 805
What Is a Markup Language? . . . 808
A Taxonomy of Markup Languages . . . 809
ASCII versus binary . . . 809
Format versus structure . . . 810
Extendable versus nonextendable . . . 811
Range of coverage . . . 812
Working with Markup . . . 812
Don't be baffled by syntax . . . 812
The language versus the interpreter . . . 814
Representing representation . . . 815
The concept of nesting . . . 815
The benefits of white space . . . 816
People play in the margins . . . 818
Summary . . . 818

Chapter 36: XML and Content Management . . . 821

What Is XML? . . . 821
XML and data interchange . . . 822
XML tagging . . . 822
Management by Schema or DTD . . . 827
Adding formatting . . . 828
Using XML in Content Management . . . 830
XML in collection . . . 830
XML in management . . . 830
XML in publishing . . . 831
XML in integration . . . 832
Help from the rest of the XML gang . . . 832

Programming in XML . . . 834
Who needs to know XML? . . . 835
Introducing DOM XML programming . . . 836
Summary . . . 843

Chapter 37: Processing Content . . . 845

What Is Content Processing? . . . 845
Stripping . . . 846
Mapping . . . 847
Content Processing and the CMS . . . 848
Focusing on the long-term benefits . . . 848
Focusing on the short-term benefits . . . 851
Distilling the essence of a process . . . 852
Bringing people into processing . . . 852
Tracking the master copy . . . 853
Managing the Processing Project . . . 855
Taking stock of the content inventory . . . 856
Drafting the processing specification . . . 856
Defining testing methods . . . 857
Getting to the Core of Content Mechanics . . . 858
Understanding the principles of mapping content . . . 860
Summary . . . 865

Chapter 38: Building Collection Systems . . . 867

Getting Started . . . 868
Minimize Disruption, Maximize Specific Value . . . 868
Forms, Files, and Batch Processes . . . 869
Driven by a Content Model . . . 870
Authoring System . . . 871
Integrated search . . . 872
Integrated workflow . . . 873
Spawning applications . . . 873
Upload support . . . 873
Advanced media support . . . 874
Spell checking . . . 874
Integrated editorial and metatorial guides . . . 875
Off-line content creation . . . 876
Preview . . . 876
Feedback . . . 877
Metadata support . . . 878
Web forms . . . 881
Other authoring applications . . . 884
Conversion System . . . 887
Basic import . . . 889
Mapping files to content types . . . 889
Batch processing . . . 890
Integration with the aggregation system . . . 890
Process management . . . 890
Quality control (QC) and monitoring . . . 892
Process and review tools . . . 894

Acquisition System . . . 896
Process and partner management . . . 897
Rights and usage . . . 897
Attribution management . . . 898
Acquiring database records . . . 898
Acquiring Web site content . . . 900
Aggregation System . . . 902
Automatically applied metadata . . . 903
Metadata standards support . . . 904
Workflow triggers . . . 904
Segmentation . . . 905
Editorial processing . . . 906
Metatorial processing . . . 909
Repository Interface . . . 914
Updates and deletes . . . 914
Submitting and storing files . . . 914
Support for WebDAV . . . 915
Summary . . . 915

Chapter 39: Building Management Systems . . . 917

What's in a Management System? . . . 918
The Repository . . . 918
Fitting into the CMS and the organization . . . 919
Getting content in and out . . . 920
A repository-wide schema . . . 920
Global search and replace . . . 921
Locating content . . . 922
Bulk processes . . . 923
Field type support . . . 923
General storage requirements . . . 924
Relational database systems . . . 926
Object databases . . . 942
File systems . . . 951
Versioning System . . . 953
Automatic versus manual versioning . . . 954
Granularity of versions . . . 954
Restoring . . . 955
Differencing versions . . . 955
Source branching . . . 956
Source Control System (Content and File Sharing) . . . 956
Sharing granularity . . . 957
Locking . . . 957
Check in and out . . . 958
Notification . . . 958
Audit trails . . . 959
Localization System . . . 959
Collection localization . . . 960
Management localization . . . 962
Publication localization . . . 966

Workflow System 968
End user interface 969
Workflow administration 972
Workflow mechanics 974
CMS Administration System 976
Administrative dashboard 977
User administration 978
Logging 978
Reporting 979
Security 980
Link support 982
Media support 982
Intellectual property tracking and costing 983
External connections 984
Robustness 987
Summary 990

Chapter 40: Building Publishing Systems 993

Templating System 994
Producing target formats 995
Producing target units structure 996
Layout and surrounds 996
Template programming 997
Producing static and dynamic publications 1000
On-the-fly conversion of text and media 1001
Integration with publication authoring tools 1002
Mix-and-match templates 1002
Navigation building 1003
Personalization System 1017
A personalization dashboard 1017
Collecting data 1018
Building rules 1021
Delivering content 1024
Deployment System 1027
Staging 1027
Content-based deployment 1028
Distributing files 1028
Scheduled publication 1029
Web System 1030
Integration with an existing Web infrastructure 1030
Searching and indexing 1032
Browser-independence 1034
Distributing files across servers 1034
Web platform support 1035
Print System 1037
Technical publications 1037
Section and subsection support 1039
Narrative support 1040
Navigation support 1044
Dynamic print publications 1045
Producing Word files 1045

E-mail System 1048
E-mail types 1049
Destination pages 1050
Integration to an e-mail server 1050
Template chooser 1051
Personalization support 1051
Syndication System 1052
Syndication subscriber management 1053
Selecting content 1054
Building and distributing feeds 1055
Other Publication Systems 1056
Multiple preview 1058
Repository Interface 1058
File and directory creation 1058
Runtime dependency resolution 1059
Database and metadata output 1059
Summary 1059

Appendix: Epilogue. 1061

Index. 1065