

Contents

Acknowledgments	vii
Introduction	xvii
Chapter 1: You: Programmer and Search Engine Marketer	1
Who Are You?	2
What Do You Need to Learn?	3
SEO and the Site Architecture	4
SEO Cannot Be an Afterthought	5
Communicating Architectural Decisions	5
Architectural Minutiae Can Make or Break You	5
Preparing Your Playground	6
Installing XAMPP	7
Preparing the Working Folder	8
Preparing the Database	11
Summary	12
Chapter 2: A Primer in Basic SEO	13
Introduction to SEO	13
Link Equity	14
Google PageRank	15
A Word on Usability and Accessibility	16
Search Engine Ranking Factors	17
On-Page Factors	17
Visible On-Page Factors	18
Invisible On-Page Factors	20
Time-Based Factors	21
External Factors	22
Potential Search Engine Penalties	26
The Google “Sandbox Effect”	26
The Expired Domain Penalty	26
Duplicate Content Penalty	27
The Google Supplemental Index	27
Resources and Tools	28
Web Analytics	28



Contents

Market Research	29
Researching Keywords	32
Browser Plugins	33
Community Forums	33
Search Engine Blogs and Resources	34
Summary	35
Chapter 3: Provocative SE-Friendly URLs	37
Why Do URLs Matter?	38
Static URLs and Dynamic URLs	38
Static URLs	39
Dynamic URLs	39
URLs and CTR	40
URLs and Duplicate Content	41
URLs of the Real World	42
Example #1: Dynamic URLs	42
Example #2: Numeric Rewritten URLs	43
Example #3: Keyword-Rich Rewritten URLs	44
Maintaining URL Consistency	44
URL Rewriting	46
Installing mod_rewrite	48
Testing mod_rewrite	49
Introducing Regular Expressions	54
URL Rewriting and PHP	60
Rewriting Numeric URLs with Two Parameters	61
Rewriting Keyword-Rich URLs	64
Building a Link Factory	66
Pagination and URL Rewriting	72
Rewriting Images and Streaming Media	72
Problems Rewriting Doesn't Solve	75
A Last Word of Caution	75
Summary	76
Chapter 4: Content Relocation and HTTP Status Codes	77
HTTP Status Codes	78
Redirection Using 301 and 302	79
301	81
302	82
Removing Deleted Pages Using 404	83
Avoiding Indexing Error Pages Using 500	84



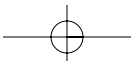
Contents

Redirecting with PHP and mod_rewrite	84
Using Redirects to Change File Names	85
URL Correction	89
Dealing with Multiple Domain Names Properly	90
Using Redirects to Change Domain Names	90
URL Canonicalization: www.example.com versus example.com	91
URL Canonicalization: /index.php versus /	92
Other Types of Redirects	94
Summary	94
Chapter 5: Duplicate Content	95
Causes and Effects of Duplicate Content	96
Duplicate Content as a Result of Site Architecture	96
Duplicate Content as a Result of Content Theft	96
Excluding Duplicate Content	97
Using the Robots Meta Tag	97
robots.txt Pattern Exclusion	99
Solutions for Commonly Duplicated Pages	103
Print-Friendly Pages	103
Navigation Links and Breadcrumb Navigation	104
Similar Pages	106
Pages with Duplicate Meta Tag or Title Values	106
URL Canonicalization	106
URL-Based Session IDs	107
Other Navigational Link Parameters	107
Affiliate Pages	108
Redirecting Parameterized Affiliate URLs	109
Summary	118
Chapter 6: SE-Friendly HTML and JavaScript	119
Overall Architecture	120
Search Engine-Friendly JavaScript	120
JavaScript Links	121
DHTML Menus	121
Popup Windows	121
DHTML Popup Windows	129
Crawlable Images and Graphical Text	129
Search Engine-Friendly HTML	140
HTML Structural Elements	141
Copy Prominence and Tables	141



Contents

Frames	144
Using Forms	144
Using a Custom Markup Language to Generate SE-Friendly HTML	145
Flash and AJAX	149
The Blended Approach	149
Summary	150
Chapter 7: Web Feeds and Social Bookmarking	151
Web Feeds	151
RSS and Atom	152
Creating RSS Feeds	154
Syndicating RSS and Atom Feeds	160
Other Sources of Syndicated Content	164
Social Bookmarking	164
Summary	172
Chapter 8: Black Hat SEO	173
What's with All the Hats?	174
Bending the Rules	175
Technical Analysis of Black-Hat Techniques	176
Attack Avoidance	177
HTML Insertion Attacks	177
Avoiding Comment Attacks Using Nofollow	180
Sanitizing User Input	184
Requesting Human Input	188
301 Redirect Attacks	194
Content Theft	196
On Buying Links	197
Digital Point Co-op, Link Vault	197
Summary	197
Chapter 9: Sitemaps	199
Traditional Sitemaps	199
Search Engine Sitemaps	200
Using Google Sitemaps	201
Using Yahoo! Sitemaps	203
Generating Sitemaps Programmatically	203
Informing Google about Updates	208



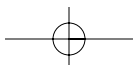
Contents

The Sitemaps.org Standard Protocol	209
Summary	210
Chapter 10: Link Bait	211
Hooking Links	211
Informational Hooks	212
News Story Hooks	212
Humor/Fun Hooks	212
Evil Hooks	212
Traditional Examples of Link Bait	213
Interactive Link Bait: Put on Your Programming Hardhat!	213
Case Study: Fortune Cookies	214
Summary	218
Chapter 11: Cloaking, Geo-Targeting, and IP Delivery	219
Cloaking, Geo-Targeting, and IP Delivery	219
More on Geo-Targeting	220
A Few Words on JavaScript Redirect Cloaking	221
The Ethical Debate on Cloaking	221
Cloaking Dangers	222
Using the Meta Noarchive Tag	222
Implementing Cloaking	223
Cloaking Case Studies	232
Rendering Images as Text	233
Redirecting Excluded Content	233
Feeding Subscription-Based Content Only to Spiders	233
Disabling URL-Based Session Handling for Spiders	234
Other Cloaking Implementations	234
Implementing Geo-Targeting	234
Summary	241
Chapter 12: Foreign Language SEO	243
Foreign Language Optimization Tips	243
Indicating Language and Region	244
Server Location and Domain Name	244
Include the Address of the Foreign Location if Possible	245
Dealing with Accented Letters (Diacritics)	245
Foreign Language Spamming	248
Summary	248



Contents

Chapter 13: Coping with Technical Issues	249
Unreliable Web Hosting or DNS	249
Changing Hosting Providers	250
Cross-Linking	251
SEO-Aware Split Testing	253
Detecting Broken Links	254
Summary	259
Chapter 14: Case Study: Building an E-Commerce Store	261
Establishing the Requirements	262
Implementing the Product Catalog	262
Summary	281
Chapter 15: Site Clinic: So You Have a Web Site?	283
1. Creating Sitemaps	284
2. Creating News Feeds	284
3. Fixing Duplication in Titles and Meta Tags	284
4. Getting Listed in Reputable Directories	284
5. Soliciting and Exchanging Relevant Links	285
6. Buying Links	285
7. Creating Link Bait	285
8. Adding Social Bookmarking Functionality	286
9. Starting a Blog and/or Forum	286
10. Dealing with a Pure Flash or AJAX Site	286
11. Preventing Black Hat Victimization	286
12. Examining Your URLs for Problems	287
13. Looking for Duplicate Content	287
14. Eliminating Session IDs	287
15. Tweaking On-page Factors	287
Summary	288
Chapter 16: WordPress: Creating an SE-Friendly Blog	289
Installing WordPress	290
Turning On Permalinks	293
Akismet: Preventing Comment Spam	294
Sociable: Social Bookmarking Icons	295
WP-Email: Email a Friend	296
Chicklet Creator Plugin	298





Contents

Sitemap Generator Plugin	299
Google Sitemaps Plugin	301
Digg Button Plugin	304
Pagerfix Plugin	305
Eliminating Duplicate Content	307
Pull-downs and Excluding Category Links	308
Excerpting Article Content	309
Making the Blog Your Home Page	309
Summary	310
Appendix A: Simple Regular Expressions	311
Matching Single Characters	312
Matching Sequences of Characters That Each Occur Once	317
Introducing Metacharacters	319
Matching Sequences of Different Characters	324
Matching Optional Characters	326
Matching Multiple Optional Characters	328
Other Cardinality Operators	332
The * Quantifier	332
The + Quantifier	334
The Curly-Brace Syntax	336
The {n} Syntax	336
The {n,m} Syntax	337
{0,m}	337
{n,m}	339
{n,}	340
Glossary	343
Index	351

