

GOOGLE HACKS



" -*kfcÄs



Tara Calishain and Rael Dornfest

O'REILLY*

Beijing • Cambridge • Farnham • Köln • Paris • Sebastopol • Taipei • Tokyo

Contents

Creditsix
Forewordxv
Prefacexvii
Chapter 1. Searching Google1
1. Setting Preferences	9
2. Language Tools	11
3. Anatomy of a Search Result	13
4. Specialized Vocabularies: Slang and Terminology	15
5. Getting Aroimd the 10 Word Limit	19
6. Word Order Matters	20
7. Repetition Matters	22
8. Mixing Syntaxes	25
9. Hacking Google URLs	29
10. Hacking Google Search Forms	31
11. Date-Range Searching	34
12. Understanding and Using Julian Dates	37
13. Using Full-Word Wildcards	37
14. inurl: Versus site:	38
15. Checking Spelling	40
16. Consulting the Dictionary	42
17. Consulting the Phonebook	43
18. Tracking Stocks	46
19. Google Interface for Translators	49

20. Searching Article Archives	53
21. Finding Directories of Information	54
22. Finding Technical Definitions	56
23. Finding Weblog Commentary	59
24. The Google Toolbar	61
25. The Mozilla Google Toolbar	63
26. The Quick Search Toolbar	64
27. GAPI	68
28. Googling with Bookmarklets	71
Chapter 2. Google Special Services and Collections.	73
29. Google Directory	76
30. Google Groups	78
31. Google Images	82
32. Google News	84
33. Google Catalogs	87
34. Froogle	88
35. Google Labs	89
Chapter 3. Third-Party Google Services.	92
36. XooML: The Google API in Plain Old XML	93
37. Google by Email	102
38. Simplifying Google Groups URLs	103
39. What Does Google Think Of...	105
40. GooglePeople	106
Chapter 4. Non-API Google Applications.	108
41. Don't Try This at Home	110
42. Building a Custom Date-Range Search Form	111
43. Building Google Directory URLs	113
44. Scraping Google Results	115
45. Scraping Google AdWords	117
46. Scraping Google Groups	121
47. Scraping Google News	125
48. Scraping Google Catalogs	128
49. Scraping the Google Phonebook	130

Chapter 5. Introducing the Google Web API.	133
50. Programming the Google Web API with Perl	142
51. Looping Around the 10-Result Limit	144
52. The SOAP::Lite Perl Module	146
53. Plain Old XML, a SOAP::Lite Alternative	150
54. NoXML, Another SOAP::Lite Alternative	154
55. Programming the Google Web API with PHP	159
56. Programming the Google Web API with Java	161
57. Programming the Google Web API with Python	163
58. Programming the Google Web API with C# and .NET	166
59. Programming the Google Web API with VB.NET	169
Chapter 6. Google Web API Applications.	173
60. Date-Range Searching with a Client-Side Application	174
61. Adding a Little Google to Your Word	178
62. Permuting a Query	179
63. Tracking Result Counts over Time	183
64. Visualizing Google Results	187
65. Meandering Your Google Neighborhood	192
66. Running a Google Popularity Contest	199
67. Building a Google Box	207
68. Capturing a Moment in Time	210
69. Feeling Really Lucky	214
70. Gleaning Phonebook Stats	217
71. Performing Proximity Searches	222
72. Blending the Google and Amazon Web Services	225
73. Getting Random Results (On Purpose)	228
74. Restricting Searches to Top-Level Results	231
75. Searching for Special Characters	236
76. Digging Deeper into Sites	238
77. Summarizing Results by Domain	241
78. Scraping Yahoo! Buzz for a Google Search	245
79. Measuring Google Mindshare	249
80. Comparing Google Results with Those of Other Search Engines	251
81. SafeSearch Certifying URLs	255

82. Syndicating Google Search Results	258
83. Searching Google Topics	259
84. Finding the Largest Page	262
85. Instant Messaging Google	265
Chapter 7. Google Pranks and Games	269
86. The No-Result Search (Prank)	269
87. Google Whacking	271
88. GooPoetry	273
89. Creating Google Art	277
90. Google Bounce	279
91. Google Mirror	282
92. Finding Recipes	284
Chapter 8. The Webmaster Side of Google	287
93. A Webmaster's Introduction to Google	289
94. Generating Google AdWords	293
95. Inside the PageRank Algorithm	294
96. 26 Steps to 15K a Day	298
97. Being a Good Search Engine Citizen	304
98. Cleaning Up for a Google Visit	307
99. Getting the Most out of AdWords	309
100. Removing Your Materials from Google	315
Index	319