

# Inhalt

|                                   |    |
|-----------------------------------|----|
| Vorwort zur dritten Auflage ..... | 13 |
| Vorwort .....                     | 15 |

## 1 Suchen im Web ..... 17

|       |                                       |    |
|-------|---------------------------------------|----|
| 1.1   | Webkataloge .....                     | 19 |
| 1.1.1 | Auswahl der Rubrik .....              | 21 |
| 1.1.2 | Titelwahl .....                       | 21 |
| 1.1.3 | Vorsicht beim Beschreibungstext ..... | 23 |
| 1.1.4 | Stichwörter mit Sorgfalt wählen ..... | 24 |
| 1.1.5 | Häufige Fehler .....                  | 24 |
| 1.1.6 | Submit-Tools .....                    | 25 |
| 1.2   | Suchmaschinen .....                   | 26 |
| 1.2.1 | User-Interface .....                  | 26 |
| 1.2.2 | Hürden .....                          | 28 |
| 1.2.3 | Funktionen und Komponenten .....      | 29 |
| 1.3   | Metasuchmaschinen .....               | 31 |
| 1.3.1 | Formale Kriterien .....               | 32 |
| 1.3.2 | Einsatzgebiet .....                   | 33 |
| 1.3.3 | Operatoren .....                      | 33 |
| 1.3.4 | Präsentation der Suchergebnisse ..... | 34 |

## 2 Anatomie des World Wide Web ..... 37

|       |                                     |    |
|-------|-------------------------------------|----|
| 2.1   | Exkurs in HTML .....                | 38 |
| 2.1.1 | HTML-Dokumentstruktur .....         | 39 |
| 2.1.2 | Tags .....                          | 40 |
| 2.1.3 | Meta-Tags .....                     | 42 |
| 2.1.4 | Sonstige Meta-Tags .....            | 49 |
| 2.1.5 | Cascading Style Sheets .....        | 51 |
| 2.2   | Trägermedium Internet .....         | 52 |
| 2.2.1 | Client-Server-Prinzip .....         | 53 |
| 2.2.2 | TCP/IP .....                        | 55 |
| 2.2.3 | Adressierung der Hosts .....        | 56 |
| 2.2.4 | Funktion und Aufbau eines URL ..... | 57 |
| 2.3   | HTTP .....                          | 59 |
| 2.3.1 | Request .....                       | 61 |

|       |                         |    |
|-------|-------------------------|----|
| 2.3.2 | Response .....          | 64 |
| 2.3.3 | HTTP live erleben ..... | 66 |

### **3 Architektur von Suchmaschinen ..... 69**

|       |                                                   |     |
|-------|---------------------------------------------------|-----|
| 3.1   | Dokumentgewinnung mit dem Webcrawler-System ..... | 70  |
| 3.1.1 | Dokumentenindex .....                             | 71  |
| 3.1.2 | Scheduler .....                                   | 72  |
| 3.1.3 | Crawler .....                                     | 74  |
| 3.1.4 | Storeserver .....                                 | 76  |
| 3.1.5 | Repository .....                                  | 81  |
| 3.2   | Datenaufbereitung und Dokumentanalyse .....       | 81  |
| 3.2.1 | Datenaufbereitung durch den Parser .....          | 85  |
| 3.2.2 | Datennormalisierung .....                         | 87  |
| 3.2.3 | Wortidentifikation durch den Tokenizer .....      | 88  |
| 3.2.4 | Identifikation der natürlichen Sprache .....      | 90  |
| 3.2.5 | Grundformreduzierung durch Word Stemming .....    | 93  |
| 3.2.6 | Mehrwortgruppenidentifikation .....               | 97  |
| 3.2.7 | Stoppwörter .....                                 | 98  |
| 3.2.8 | Keyword-Extrahierung .....                        | 99  |
| 3.2.9 | URL-Verarbeitung .....                            | 104 |
| 3.3   | Datenstruktur .....                               | 104 |
| 3.3.1 | Hitlist .....                                     | 105 |
| 3.3.2 | Direkter Index .....                              | 108 |
| 3.3.3 | Invertierter Index .....                          | 110 |
| 3.3.4 | Verteilte Datenstruktur .....                     | 111 |

### **4 Gewichtung und Relevanz ..... 115**

|       |                                                       |     |
|-------|-------------------------------------------------------|-----|
| 4.1   | Statistische Modelle .....                            | 117 |
| 4.1.1 | Boolesches Retrieval .....                            | 117 |
| 4.1.2 | Fuzzy-Logik .....                                     | 118 |
| 4.1.3 | Vektorraummodell .....                                | 119 |
| 4.1.4 | Relative Worthäufigkeit (TF) .....                    | 122 |
| 4.1.5 | Inverse Dokumenthäufigkeit (IDF) .....                | 123 |
| 4.1.6 | Bedeutung der Lage und Auszeichnung eines Terms ..... | 124 |
| 4.1.7 | Betrachtung des URL .....                             | 125 |
| 4.2   | Pagerank .....                                        | 125 |
| 4.2.1 | Link-Popularity .....                                 | 126 |
| 4.2.2 | Pagerank-Konzept und Random Surfer .....              | 127 |

|       |                                                        |     |
|-------|--------------------------------------------------------|-----|
| 4.2.3 | Pagerank-Formel .....                                  | 128 |
| 4.2.4 | Beispiel zur Pagerank-Berechnung .....                 | 129 |
| 4.2.5 | Effekte des Pageranks .....                            | 131 |
| 4.2.6 | Intelligente Surfer und weitere Einflussfaktoren ..... | 132 |
| 4.2.7 | Bad-Rank .....                                         | 135 |
| 4.3   | Click-Popularity .....                                 | 137 |
| 4.4   | Cluster-Verfahren .....                                | 139 |
| 4.4.1 | Cluster-Verfahren im Einsatz .....                     | 140 |
| 4.4.2 | Vivisimo – ein Pionier .....                           | 141 |
| 4.4.3 | Single-Pass-Methode .....                              | 143 |

## **5 Suchprozess ..... 145**

|       |                                            |     |
|-------|--------------------------------------------|-----|
| 5.1   | Arbeitsschritte des Query-Prozessors ..... | 146 |
| 5.1.1 | Tokenizing .....                           | 146 |
| 5.1.2 | Parsing .....                              | 146 |
| 5.1.3 | Stoppwörter und Stemming .....             | 147 |
| 5.1.4 | Erzeugung der Query .....                  | 147 |
| 5.1.5 | Verwendung eines Thesaurus .....           | 148 |
| 5.1.6 | Matching und Gewichtung .....              | 148 |
| 5.1.7 | Darstellung der Trefferliste .....         | 149 |
| 5.2   | Suchoperatoren .....                       | 150 |
| 5.2.1 | Boolesche Ausdrücke .....                  | 150 |
| 5.2.2 | Phrasen .....                              | 151 |
| 5.2.3 | Wortabstand .....                          | 152 |
| 5.2.4 | Trunkierung .....                          | 153 |
| 5.3   | Erweiterte Suchmöglichkeiten .....         | 153 |
| 5.3.1 | Sprachfilter .....                         | 155 |
| 5.3.2 | Positionierung .....                       | 155 |
| 5.3.3 | Aktualität .....                           | 156 |
| 5.3.4 | Domainfilter .....                         | 156 |
| 5.3.5 | Dateityp .....                             | 157 |
| 5.3.6 | Sonstige Suchmöglichkeiten .....           | 157 |
| 5.4   | Nutzerverhalten im Web .....               | 159 |
| 5.4.1 | Suchaktivitäten .....                      | 159 |
| 5.4.2 | Suchmodi .....                             | 161 |
| 5.4.3 | Welche Suchmaschine wird genutzt? .....    | 163 |
| 5.4.4 | Was wird gesucht? .....                    | 165 |

|          |                                                        |            |
|----------|--------------------------------------------------------|------------|
| <b>6</b> | <b>Keyword-Recherche .....</b>                         | <b>169</b> |
| 6.1      | Ausrichtung einer Website .....                        | 169        |
| 6.1.1    | Zielgruppe .....                                       | 170        |
| 6.1.2    | Zielsetzung .....                                      | 171        |
| 6.1.3    | Leads und Konversionsraten .....                       | 172        |
| 6.2      | Gütekriterien .....                                    | 173        |
| 6.3      | Erstellen einer Keyword-Liste .....                    | 176        |
| 6.3.1    | Erstes Brainstorming .....                             | 177        |
| 6.3.2    | Logfiles nutzen .....                                  | 178        |
| 6.3.3    | Mitbewerber analysieren .....                          | 179        |
| 6.3.4    | Synonyme finden .....                                  | 181        |
| 6.3.5    | Umfeld: Freunde, Kollegen, Bekannte und Besucher ..... | 182        |
| 6.3.6    | IDF überprüfen .....                                   | 184        |
| 6.3.7    | Erste Bereinigung .....                                | 185        |
| 6.3.8    | Keyword-Datenbanken .....                              | 186        |
| 6.4      | Eigenschaften der Keywords .....                       | 193        |
| 6.4.1    | Groß- und Kleinschreibung .....                        | 193        |
| 6.4.2    | Singular oder Plural .....                             | 194        |
| 6.4.3    | Sonderzeichen .....                                    | 194        |
| 6.4.4    | Sonstige Eigenschaften .....                           | 195        |
| 6.4.5    | Falsche orthografische Schreibweise .....              | 195        |
| 6.4.6    | Getrennt oder zusammen? .....                          | 196        |
| 6.4.7    | Wortkombinationen und Wortnähe .....                   | 197        |
| 6.5      | Bewerten der Listeneinträge .....                      | 199        |
| 6.5.1    | Liste bereinigen .....                                 | 199        |
| 6.5.2    | Keyword-Effizienz berechnen .....                      | 200        |
| 6.5.3    | Finale Auswahl und Zuweisung von Seiten-Keywords ....  | 204        |
| <b>7</b> | <b>Onpage-Optimierung .....</b>                        | <b>207</b> |
| 7.1      | Spezielle Situation bei einem Relaunch .....           | 209        |
| 7.2      | Strukturelle Vorbereitungen .....                      | 211        |
| 7.2.1    | Gültiges HTML .....                                    | 211        |
| 7.2.2    | Einsatz von CSS .....                                  | 213        |
| 7.2.3    | Seitenstruktur .....                                   | 214        |
| 7.2.4    | Navigation .....                                       | 218        |
| 7.2.5    | Frames .....                                           | 221        |
| 7.2.6    | Die Startseite .....                                   | 226        |
| 7.2.7    | Dateityp und dynamische Seiten .....                   | 228        |

|        |                                                                                 |     |
|--------|---------------------------------------------------------------------------------|-----|
| 7.3    | Optimierung durch Tags .....                                                    | 234 |
| 7.3.1  | Title .....                                                                     | 234 |
| 7.3.2  | Fließtext und Keyword-Dichte .....                                              | 237 |
| 7.3.3  | Aufzählungen .....                                                              | 240 |
| 7.3.4  | Texthervorhebungen .....                                                        | 240 |
| 7.3.5  | Überschriften .....                                                             | 242 |
| 7.3.6  | Links und Anchor-Text .....                                                     | 244 |
| 7.3.7  | Tabellen .....                                                                  | 246 |
| 7.3.8  | Bilder und Image-Maps .....                                                     | 248 |
| 7.3.9  | Phantom-Pixel .....                                                             | 249 |
| 7.3.10 | Comment .....                                                                   | 250 |
| 7.3.11 | Form und Input .....                                                            | 251 |
| 7.3.12 | Noscript .....                                                                  | 252 |
| 7.3.13 | Iframe .....                                                                    | 253 |
| 7.4    | Web 2.0 und AJAX für die Onpage-Optimierung .....                               | 254 |
| 7.4.1  | AJAX kurz vorgestellt .....                                                     | 254 |
| 7.4.2  | Sorgenkind AJAX bei der Onpage-Optimierung .....                                | 256 |
| 7.4.3  | Richtlinien für den Einsatz von AJAX für die<br>Suchmaschinen-Optimierung ..... | 257 |
| 7.5    | PDF-Dokumente optimieren .....                                                  | 258 |

## **8 Offpage-Optimierung ..... 261**

|       |                                                   |     |
|-------|---------------------------------------------------|-----|
| 8.1   | Webserver und Restriktionen .....                 | 261 |
| 8.1.1 | Webhosting .....                                  | 261 |
| 8.1.2 | Restriktionen .....                               | 263 |
| 8.2   | Domainname und Verzeichnisse .....                | 264 |
| 8.2.1 | Domainname .....                                  | 264 |
| 8.2.2 | Verzeichnis- und Dateinamen .....                 | 266 |
| 8.2.3 | Verzeichnistiefe und Aktualität .....             | 268 |
| 8.3   | Sitestructur .....                                | 271 |
| 8.3.1 | Redirects korrekt umsetzen .....                  | 272 |
| 8.3.2 | Deep Web .....                                    | 274 |
| 8.3.3 | Seiten ausschließen (robots.txt) .....            | 276 |
| 8.4   | Link-Popularity erhöhen .....                     | 279 |
| 8.4.1 | Interne Verlinkung optimieren .....               | 279 |
| 8.4.2 | Das KAKADU-Prinzip .....                          | 280 |
| 8.4.3 | Qualitätskriterien potenzieller Linkpartner ..... | 283 |
| 8.4.4 | An andere Webautoren herantreten .....            | 284 |
| 8.4.5 | Eingehende Links erzielen .....                   | 285 |

|       |                                              |     |
|-------|----------------------------------------------|-----|
| 8.4.6 | Link-Farmen und Google-Bomben .....          | 289 |
| 8.4.7 | Aufbau von Satelliten-Domains .....          | 290 |
| 8.5   | Web 2.0 zur Offpage-Optimierung nutzen ..... | 293 |
| 8.5.1 | Wikis nutzen .....                           | 293 |
| 8.5.2 | Social Bookmarking .....                     | 296 |
| 8.5.3 | Nofollow-Follow-These .....                  | 297 |
| 8.5.4 | Web 2.0-Nutzer arbeiten lassen .....         | 299 |
| 8.5.5 | RSS Feeds anbieten .....                     | 300 |
| 8.6   | Click-Popularity erhöhen .....               | 302 |

## **9 Spam ..... 305**

|      |                                    |     |
|------|------------------------------------|-----|
| 9.1  | Keyword-Stuffing .....             | 307 |
| 9.2  | Unsichtbare und kleine Texte ..... | 308 |
| 9.3  | Hidden-Links .....                 | 314 |
| 9.4  | Meta-Spam .....                    | 315 |
| 9.5  | Doorway-Pages .....                | 316 |
| 9.6  | Cloaking .....                     | 320 |
| 9.7  | Bait-And-Switch .....              | 323 |
| 9.8  | Domain-Dubletten .....             | 324 |
| 9.9  | Page-Jacking .....                 | 326 |
| 9.10 | Blog- und Gästebuch-Spam .....     | 327 |
| 9.11 | Sonstige Spam-Methoden .....       | 329 |

## **10 Aufnahme in die Suchmaschinen ..... 331**

|        |                                   |     |
|--------|-----------------------------------|-----|
| 10.1   | Suchmaschinen-Kooperationen ..... | 331 |
| 10.2   | Die Anmeldung .....               | 333 |
| 10.2.1 | Manuelle Anmeldung .....          | 335 |
| 10.2.2 | Automatische Anmeldung .....      | 338 |
| 10.2.3 | Aufnahmedauer .....               | 339 |
| 10.3   | Kostenpflichtige Leistungen ..... | 341 |
| 10.3.1 | Payed-Inclusion-Programme .....   | 342 |
| 10.3.2 | Pay-Per-Click (PPC) .....         | 344 |

## **11 Monitoring und Controlling ..... 349**

|        |                                  |     |
|--------|----------------------------------|-----|
| 11.1   | Server-Monitoring .....          | 350 |
| 11.2   | Controlling .....                | 353 |
| 11.3   | Logfile-Analyse .....            | 355 |
| 11.3.1 | Anfragen pro Tag und Monat ..... | 357 |

|        |                                            |     |
|--------|--------------------------------------------|-----|
| 11.3.2 | Herkunftsland der Besucher .....           | 359 |
| 11.3.3 | Seitenbesuche .....                        | 359 |
| 11.3.4 | Herkunft der Besucher .....                | 360 |
| 11.3.5 | Besuche über Suchmaschinen .....           | 361 |
| 11.3.6 | Suchbegriffe .....                         | 362 |
| 11.3.7 | Sonstige Informationen .....               | 363 |
| 11.4   | Rank-Monitoring .....                      | 364 |
| 11.5   | Einträge aus Suchmaschinen entfernen ..... | 367 |

## **12 Google – Gerüchte, Theorien und Fakten ..... 369**

|        |                                          |     |
|--------|------------------------------------------|-----|
| 12.1   | Gerüchtequellen und Gerüchteküchen ..... | 369 |
| 12.2   | Googles Crawling-Strategien .....        | 370 |
| 12.2.1 | Everflux .....                           | 370 |
| 12.2.2 | Fresh Crawl und Deep Crawl .....         | 371 |
| 12.3   | Die Google-Updates .....                 | 371 |
| 12.3.1 | Varianten von Updates .....              | 372 |
| 12.3.2 | Update Historie .....                    | 373 |
| 12.4   | Google und die geheimen Labors .....     | 379 |
| 12.4.1 | Geheime Labors .....                     | 380 |
| 12.4.2 | Trustcenter .....                        | 380 |
| 12.5   | Sandbox .....                            | 381 |
| 12.5.1 | Der Sandbox-Effekt .....                 | 381 |
| 12.5.2 | Gerücht oder Fakt? .....                 | 382 |
| 12.5.3 | Den Sandbox-Effekt vermeiden .....       | 382 |
| 12.6   | Hilltop-Prinzip und Trustrank .....      | 385 |
| 12.7   | Google-Sitemap-Programm .....            | 387 |

## **13 Usability und Suchmaschinen-Optimierung ..... 391**

|        |                                                              |     |
|--------|--------------------------------------------------------------|-----|
| 13.1   | Suchmaschinen-Optimierung alleine reicht nicht .....         | 392 |
| 13.2   | Was Usability mit Suchmaschinen-Optimierung zu tun hat ..... | 393 |
| 13.2.1 | Was ist Usability? .....                                     | 393 |
| 13.2.2 | Von der Suchmaschinen-Optimierung zur Usability .....        | 395 |
| 13.3   | Usability-Regeln .....                                       | 399 |
| 13.3.1 | Kohärenz und Konsistenz .....                                | 399 |
| 13.3.2 | Erwartungen erfüllen .....                                   | 400 |
| 13.3.3 | Schnelle Erschließbarkeit .....                              | 402 |
| 13.3.4 | Lesbarkeit sicherstellen .....                               | 403 |
| 13.3.5 | Nutzersicht einnehmen! .....                                 | 405 |
| 13.3.6 | Zweckdienliche und einfache Navigation .....                 | 406 |

## 14 Optimierung konkret: TYPO3 und WordPress ..... 409

|         |                                                    |     |
|---------|----------------------------------------------------|-----|
| 14.1    | CMS optimieren am Beispiel von TYPO3 .....         | 410 |
| 14.1.1  | Vorbereitungen zur Optimierung .....               | 410 |
| 14.1.2  | Suchmaschinenfreundliche URL mit AliasPro .....    | 413 |
| 14.1.3  | Noch besser: RealURL .....                         | 413 |
| 14.1.4  | Title-Tag in TYPO3 .....                           | 414 |
| 14.1.5  | Meta-Tags automatisch setzen .....                 | 416 |
| 14.1.6  | Breadcrumb-Navigation einbinden .....              | 417 |
| 14.1.7  | Sitemap erstellen .....                            | 418 |
| 14.1.8  | Google-Sitemap einbinden .....                     | 419 |
| 14.2    | Weblogs optimieren am Beispiel von WordPress ..... | 420 |
| 14.2.1  | Bloggen und Suchmaschinen-Optimierung .....        | 420 |
| 14.2.2  | Schreiben für Leser und Suchmaschinen .....        | 421 |
| 14.2.3  | Suchmaschinenfreundliches Template .....           | 422 |
| 14.2.4  | Title-Tag .....                                    | 423 |
| 14.2.5  | Überschriften und Textauszeichnungen .....         | 425 |
| 14.2.6  | Blog-URLs optimieren .....                         | 426 |
| 14.2.7  | Plugins als URL-Helferchen .....                   | 427 |
| 14.2.8  | Crawler im Geschwindigkeitswahn .....              | 428 |
| 14.2.9  | Kommentare auslagern .....                         | 428 |
| 14.2.10 | Google-Sitemap in WordPress erzeugen .....         | 429 |
| 14.2.11 | Interne Verlinkung stärken .....                   | 430 |
| 14.2.12 | Ansätze zur Offpage-Optimierung .....              | 430 |
| 14.2.13 | Nofollow deaktivieren .....                        | 432 |
| 14.2.14 | Content is King .....                              | 433 |

## Anhang ..... 435

|   |                             |     |
|---|-----------------------------|-----|
| A | Glossar .....               | 437 |
| B | Literaturverzeichnis .....  | 451 |
| C | Quellen .....               | 453 |
| D | Abbildungsverzeichnis ..... | 459 |
|   | Index .....                 | 463 |