

Dirk Lewandowski

Suchmaschinen verstehen

3. Auflage

 Springer Vieweg

Suchmaschinen verstehen

Dirk Lewandowski

Suchmaschinen verstehen

3., vollständig überarbeitete und erweiterte Auflage

Dirk Lewandowski
Hochschule für Angewandte Wissenschaften Hamburg
Hamburg, Deutschland

ISBN 978-3-662-63190-4 ISBN 978-3-662-63191-1 (eBook)
<https://doi.org/10.1007/978-3-662-63191-1>

Die Deutsche Nationalbibliothek verzeichnet diese Publikation in der Deutschen Nationalbibliografie; detaillierte bibliografische Daten sind im Internet über <http://dnb.d-nb.de> abrufbar.

Springer Vieweg

© Springer-Verlag GmbH Deutschland, ein Teil von Springer Nature 2015, 2018, 2021

Das Werk einschließlich aller seiner Teile ist urheberrechtlich geschützt. Jede Verwertung, die nicht ausdrücklich vom Urheberrechtsgesetz zugelassen ist, bedarf der vorherigen Zustimmung des Verlags. Das gilt insbesondere für Vervielfältigungen, Bearbeitungen, Übersetzungen, Mikroverfilmungen und die Einspeicherung und Verarbeitung in elektronischen Systemen.

Die Wiedergabe von allgemein beschreibenden Bezeichnungen, Marken, Unternehmensnamen etc. in diesem Werk bedeutet nicht, dass diese frei durch jedermann benutzt werden dürfen. Die Berechtigung zur Benutzung unterliegt, auch ohne gesonderten Hinweis hierzu, den Regeln des Markenrechts. Die Rechte des jeweiligen Zeicheninhabers sind zu beachten.

Der Verlag, die Autoren und die Herausgeber gehen davon aus, dass die Angaben und Informationen in diesem Werk zum Zeitpunkt der Veröffentlichung vollständig und korrekt sind. Weder der Verlag, noch die Autoren oder die Herausgeber übernehmen, ausdrücklich oder implizit, Gewähr für den Inhalt des Werkes, etwaige Fehler oder Äußerungen. Der Verlag bleibt im Hinblick auf geografische Zuordnungen und Gebietsbezeichnungen in veröffentlichten Karten und Institutionsadressen neutral.

Planung: Petra Steinmüller

Springer Vieweg ist ein Imprint der eingetragenen Gesellschaft Springer-Verlag GmbH, DE und ist ein Teil von Springer Nature.

Die Anschrift der Gesellschaft ist: Heidelberger Platz 3, 14197 Berlin, Germany

Vorwort zur dritten Auflage

Es ist mir eine Freude, dass dieses Buch nun bereits in der dritten Auflage vorliegt. Die zahlreichen Rückmeldungen von Leserinnen und Lesern der vergangenen Auflagen haben mir nicht nur gezeigt, dass ein Buch, welches einen Überblick über alle wesentlichen Facetten des Themas Suchmaschinen und Web-Suche gibt, nötig ist. Vielmehr habe ich durch die vielen Rückmeldungen auch zahlreiche Hinweise bekommen, welche Teile noch verständlicher gemacht werden sollten, wo ergänzt, und wo gestrichen werden sollte.

Diese neue Auflage wurde aber nicht nur in den Details überarbeitet, sondern inhaltlich etwas umstrukturiert, um den aktuellen Gegebenheiten gerecht zu werden: So wurde das Kapitel zu den Spezialsuchmaschinen bzw. zu den speziellen Kollektionen, die in die Web-Suche eingehen, gemäß der immer weiter zunehmenden Bedeutung dieser Kollektionen als nun sechstes Kapitel verschoben und vollständig neu strukturiert und überarbeitet. Das Kapitel knüpft nun direkt an das Ranking der allgemeinen Web-Inhalte an und führt auf das folgende Kapitel zur Präsentation der Suchergebnisse hin. Das Kapitel zu Social Media wurde gekürzt und in das Kapitel zum Deep Web integriert. Zwar ist die Bedeutung von Social Media bei weitem nicht zurückgegangen, allerdings hat die noch vor einigen Jahren von vieler Seite angenommene, starke Einbindung der bekannten Social-Media-Dienste in die Suchmaschinen nicht stattgefunden. Daher handelt es sich bei den Inhalten dieser Dienste nun (weiterhin) um Inhalte des Deep Web, die allerdings aufgrund ihrer Bedeutung weiterhin in einem eigenen Abschnitt behandelt werden.

Die Literatur zum Thema Suchmaschinen ist mittlerweile fast unüberschaubar; vieles ist allerdings zu spezialisiert, um in ein Überblickswerk aufgenommen zu werden. Die Verweise auf die Literatur wurden vielfach erheblich ergänzt; ich habe aber darauf geachtet, dass diesem Buch nicht das Schicksal vieler einführender Werke widerfährt, nämlich, dass es immer weiter anwächst, letztlich aber keine gut lesbare Einführung mehr darstellt.

Alle im Text genannten Zahlen, beispielsweise zur Internet- und Suchmaschinennutzung, wurden anhand der neuesten Studien und Erhebungen auf den aktuellen Stand gebracht. Auch die Screenshots der Suchergebnisseiten wurden durchgehend aktualisiert. Teilweise haben sich in der Ergebnisdarstellung nur geringfügigen Änderungen ergeben; der Anspruch war allerdings, dass das Buch durchgehend den aktuellen Stand zum

Zeitpunkt der Manuskriptabgabe zeigt. Ebenso wurden alle im Text genannten URLs überprüft und auf den aktuellen Stand gebracht.

Und nicht zuletzt wurde der Text hin zu einer gendergerechten Sprache angepasst. Leider gibt es noch keine Regeln für solche Anpassungen, die auch zu einem sprachlich befriedigenden Text führen. Daher wurde in manchen Fällen das generische Maskulinum belassen.

Mein besonderer Dank gilt allen, die die Überarbeitung des Manuskripts durch ihre Anmerkungen zu den vergangenen Auflagen und durch zahlreiche Diskussionen unterstützt haben. Für die kontinuierliche Diskussion über die letzten Jahre hinweg und die Freude, die sie mir bei meiner Tätigkeit an der HAW Hamburg geben, bedanke ich mich bei dem kleinen Search-Studies-Team, vor allem bei Sebastian Sünkler und Sebastian Schultheiß. Bei der Überarbeitung des Manuskripts haben auch wieder studentische Hilfskräfte mitgeholfen. Bei Nina Niesche bedanke ich mich für die Überarbeitung der Screenshots und zahlreicher Schaubilder; bei Ngan Pham für die Überarbeitung der Literaturverzeichnisse. Nicht vergessen werden dürfen natürlich auch diejenigen, die bei der Erstellung der ersten Auflagen maßgeblich geholfen haben und deren Arbeit sich natürlich auch in der aktuellen Version wiederfindet: Dorothee Wagner, Alexandra Linhart und Jenny Peters. Vielen Dank!

Hamburg, Deutschland
Februar 2021

Dirk Lewandowski

Inhaltsverzeichnis

1 Einführung	1
1.1 Die Bedeutung der Suchmaschinen	2
1.2 Ein Buch über Google?	6
1.3 Ziel dieses Buchs	7
1.4 Über Suchmaschinen sprechen	8
1.5 Aufbau des Buchs	8
1.6 Aufbau der Kapitel und Markierungen innerhalb des Texts	11
1.7 Zusammenfassung	11
Literatur	11
2 Formen der Suche im Web	13
2.1 Suche nach einer Website vs. Suche nach Informationen zu einem Thema	13
2.2 Was ist ein Dokument?	15
2.3 Wo wird gesucht?	16
2.4 Zugänge zu den Informationen im World Wide Web	16
2.4.1 Suchmaschinen	17
2.4.2 Spezialsuchmaschinen	20
2.4.3 Metasuchmaschinen	22
2.4.4 Web-Verzeichnisse	23
2.4.5 Social-Bookmarking-Dienste	25
2.4.6 Frage-Antwort-Dienste	25
2.4.7 Soziale Netzwerke	26
2.5 Zusammenfassung	26
Literatur	27
3 Wie Suchmaschinen die Inhalte des Web erfassen und aufbereiten	29
3.1 Das World Wide Web und seine Erfassung durch Suchmaschinen	33
3.2 Content Acquisition	36
3.3 Web Crawling: Dokumente im Web aufspüren	38
3.3.1 Suchmaschinen steuern und ausschließen	42
3.3.2 Ausschluss von Inhalten durch die Suchmaschinenbetreiber	45

3.3.3	Aufbau der Datenbasis und Crawling für spezielle Kollektionen	47
3.4	Der Indexer: Dokumente für die Suche aufbereiten	48
3.4.1	Indexierung von Bildern, Audiodateien und Videos	54
3.4.2	Repräsentation von Web-Dokumenten in Suchmaschinen	55
3.5	Der Searcher: Suchanfragen verstehen	58
3.6	Zusammenfassung	62
	Literatur	63
4	Wie Suchmaschinen genutzt werden	67
4.1	Der Suchprozess	67
4.2	Erfassung von Nutzungsdaten	69
4.3	Anfragetypen	70
4.4	Sessions	73
4.5	Suchanfragen	75
4.5.1	Eingabe der Suchanfragen	75
4.5.2	Suchvorschläge während der Eingabe	78
4.5.3	Formulierung der Suchanfragen	79
4.5.4	Länge der Suchanfragen	81
4.5.5	Verteilung der Suchanfragen nach Häufigkeiten	83
4.5.6	Suchanfragen-Trends	86
4.5.7	Verwendung von Operatoren und Befehlen zur gezielten Suche. . .	87
4.6	Themen	88
4.7	Zusammenfassung	89
	Literatur	90
5	Das Ranking der Suchergebnisse	93
5.1	Gruppen von Rankingfaktoren	95
5.2	Textstatistik	96
5.2.1	Ermittlung potenziell relevanter Dokumente	97
5.2.2	Berechnung von Häufigkeiten	99
5.2.3	Berücksichtigung von Strukturinformationen in Dokumenten . . .	101
5.3	Popularität	103
5.3.1	Linktopologische Verfahren	104
5.3.2	Nutzungsstatistische Verfahren	109
5.4	Aktualität	117
5.5	Lokalität	119
5.6	Personalisierung der Suchergebnisse	124
5.7	Technische Rankingfaktoren	127
5.8	Ranking und Spam	128
5.9	Zusammenfassung	130
	Literatur	131

6 Vertikale Suche: Spezielle Kollektionen und ihre Einbindung in die allgemeine Suche	135
6.1 Spezialsuchmaschinen als Basis der Universal Search	137
6.2 Typen von Spezialsuchmaschinen	140
6.3 Beispiele für Kollektionen	142
6.3.1 Nachrichten	143
6.3.2 Wissenschaftliche Inhalte	147
6.3.3 Bilder	150
6.3.4 Videos	151
6.4 Einbindung von Spezialsuchmaschinen in die Universal Search	152
6.5 Zusammenfassung	153
Literatur	154
7 Die Präsentation der Suchergebnisse	157
7.1 Der Einfluss von Endgerät und Bildschirmauflösung	158
7.2 Der Aufbau der Suchergebnisseiten	159
7.3 Elemente auf den Suchergebnisseiten	168
7.3.1 Organische Ergebnisse	168
7.3.2 Werbung	169
7.3.3 Universal-Search-Ergebnisse	170
7.3.4 Knowledge-Graph-Ergebnisse	172
7.3.5 Direkte Antworten	172
7.3.6 Einbindung von Transaktionen auf der Suchergebnisseite	175
7.3.7 Navigationselemente	175
7.3.8 Hilfestellungen zur Veränderung der Suchanfrage	175
7.3.9 Suchoptionen auf der Ergebnisseite	178
7.4 Der Aufbau der Trefferbeschreibungen	178
7.5 Optionen zu den einzelnen Ergebnissen	182
7.6 Auswahl von geeigneten Treffern auf den Suchergebnisseiten	184
7.7 Zusammenfassung	185
Literatur	186
8 Der Suchmaschinenmarkt	189
8.1 Das Geschäftsmodell der Suchmaschinen	189
8.2 Die Bedeutung der Suchmaschinen für die Online-Werbung	191
8.3 Marktanteile der Suchmaschinen	192
8.4 Bedeutende Suchmaschinen	194
8.5 Partnerschaften auf dem Suchmaschinenmarkt	195
8.6 Zusammenfassung	197
Literatur	198
9 Suchmaschinenoptimierung	201
9.1 Die Bedeutung der Suchmaschinenoptimierung	202

9.2	Grundlegende Verfahren der Suchmaschinenoptimierung.	204
9.2.1	Inhalte	207
9.2.2	Website-Architektur.	208
9.2.3	HTML	209
9.2.4	Vertrauen	209
9.2.5	Links	210
9.2.6	Personenbezogene Faktoren.	210
9.2.7	Schädliche Faktoren.	211
9.2.8	Spezielsuchen.	212
9.3	Suchmaschinenoptimierung und Spam	212
9.4	Die Bedeutung von Rankingupdates	213
9.5	Suchmaschinenoptimierung für besondere Kollektionen	214
9.6	Die Position der Suchmaschinenbetreiber	215
9.7	Zusammenfassung	216
	Literatur.	217
10	Suchmaschinenwerbung.	219
10.1	Besonderheiten der Suchmaschinenwerbung.	222
10.2	Funktionsweise und Ranking	224
10.3	Unterscheidbarkeit von Werbung und organischen Suchergebnissen	226
10.4	Werbung innerhalb der Universal-Search-Ergebnisse	227
10.5	Zusammenfassung	228
	Literatur.	229
11	Alternativen zu Google.	231
11.1	Überschneidungen zwischen den Suchergebnissen verschiedener Suchmaschinen	232
11.2	Warum sollte man eine andere Suchmaschine als Google nutzen?	233
11.2.1	Einholen einer „zweiten Meinung“	233
11.2.2	Mehr bzw. zusätzliche Ergebnisse	234
11.2.3	Andere Ergebnisse	235
11.2.4	Bessere Ergebnisse.	236
11.2.5	Andere Trefferpräsentation	236
11.2.6	Andere Benutzerführung	236
11.2.7	Verhinderung der Bildung von Nutzerprofilen	237
11.2.8	Andere Suchmöglichkeiten	237
11.3	Wann sollte man eine andere Suchmaschine als Google nutzen?	237
11.4	Besonderheiten bei Google aufgrund seiner Marktdominanz	239
11.5	Zusammenfassung	242
	Literatur.	243