

Information Retrieval und Automatisierung Perspektiven und Probleme

Andrew E. Wessel

Technische Hochschule Darmstadt
FACHBEREICH INFORMATIK
B I B L I O T H E K
Inventar-Nr.: 2547
Sachgebiete: _____
Standort: _____

Luchterhand

	Vorwort der Herausgeber	10
	Vorwort des Verfassers	11
	Anmerkung zur deutschen Ausgabe	12
	Prolog – Zukunft und Vergangenheit	13
1	Gut strukturierte und schlecht strukturierte Informationen	16
1.1	Das Problem der Analyse und Indexierung	20
2	Indexierung und Analyse von Informationen – Einige einleitende Begriffe	22
2.1	Freitext-Indexierung (free text indexing)	23
2.2	Thesauri-Indexierung (thesauri indexing)	24
3	Freitext-Indexierung – Die Probleme und Paradoxia	27
3.1	Automatisierung und Textanalyse – Die Hoffnung auf eine allgemeine Lösung	30
3.2	Formale Erwägungen – Theoretische Unvollständigkeit und Widersprüche	32
4	Frühere Ansätze und Ergebnisse – Eine Geschichte von zwei Kontinenten	36
4.1	Das Europäische Dokumentationszentrum für Weltraumforschung (ESRO/ELDO-RECON)	38
4.2	Patentdokumentation	43
4.2.1	Probleme der automatisierten Indexierung für die Klassifikation und die Feinindexierung von Patenten	48
4.2.1.1	Feinindexierung und Klassifikation – Einige Probleme der Experimentanlage	51
4.2.2	Projekt Potomac – Das Klassifikationsexperiment des U.S. Patent Office	55
5	Automatisierung gegen Teilautomatisierung	59
5.1	Frühe Experimente zur Durchführbarkeit der teilautomatischen Indexierung – Geschichte und Ergebnisse	61
5.2	Vernünftige und andere Gründe	65

8	6	Thesauri, Computer und Menschen – Ein realisierbarer Ansatz	69
	6.1	Die Probleme und Prozesse	70
	6.2	Computerhilfen für die Entwicklung und den Einsatz von Thesauri	75
	6.3	Allgemeine Kostenerwägungen	81
	7	Grundlegende Prinzipien der Erstellung von Thesauri für das automatisierte Information Retrieval	84
	7.1	Thesauri-Deskriptoren	85
	7.1.1	Inhaltsdeskriptoren (content descriptors)	85
	7.1.2	Dokumentdeskriptoren (document descriptors)	89
	7.1.3	Metadeskriptoren (meta-descriptors)	90
	7.1.4	Nähere Bestimmungen zu Deskriptoren (descriptor qualifiers)	92
	7.1.5	Bereichsindikatoren (scope indicators)	94
	7.2	Erläuterungen (explications)	95
	7.3	Interpretationsregeln (interpretive rules)	97
	7.4	Anwendungsregeln (applicative rules)	99
	7.4.1	Indexierungsregeln (indexing rules)	99
	7.4.2	Rechercheregeln (search rules)	102
	7.5	Eine Anmerkung zu ergänzenden Computerhilfen	104
	8	Erläuterung der grundlegenden Prinzipien der Erstellung von Thesauri für das automatisierte Information Retrieval	105
	8.1	Thesaurus der Stiftung Wissenschaft und Politik (SWP-T) – Gesamtüberblick	105
	8.2	Deskriptoren	107
	8.2.1	Dokumentdeskriptoren	107
	8.2.1.1	Nähere Bestimmungen zu Dokumentdeskriptoren (document descriptor qualifiers)	110
	8.2.1.2	Namenklassen-Dokumentdeskriptoren (Name Class document descriptors)	112
	8.2.2	Inhaltsdeskriptoren	115
	8.2.2.1	Inhaltsdeskriptoren im SWP-T	117
	8.2.2.1.1	Internationale Politik	117
	8.2.3	Erläuterungen und Interpretationsregeln in Thesauri	123
	8.2.3.1	Erläuterungen	124
	8.2.3.2	Interpretationsregeln	131
	8.2.4	Anwendungsregeln	139
	8.2.4.1	Regeln für die Eingabe in den Computer	139
	8.2.4.2	Obligatorische Deskriptorbeziehungen und Ergänzungsformate	141

9	Computer-Software und einige Hardware-Erwägungen	146
9.1	Probleme der Dateistrukturierung	149
9.1.1	Computerdateien – Ältere Formen für die Stapelverarbeitung	151
9.1.2	Computerdateien – Moderne Dialogsysteme	152
9.2	Spezifische Computer-Software-Probleme für Information Retrieval Systeme im Dialogbetrieb	154
9.2.1	Bits and Bytes	157
9.2.2	Formatgröße, Deskriptoren und Dokumente	158
9.2.3	Ein sinnvoller Kompromiß bei fester Satzlänge	163
9.2.3.1	Reihenfolge der Adressen und zurück zu Bits und Bytes	166
9.2.4	Dialogdateien	171
9.2.5	Dialogfunktionen	176
9.2.5.1	Mehrfacheingabe und Fehler bei der Dokumentidentifizierung	177
9.2.5.2	Betriebssysteme für mehrere Benutzer und Probleme der Änderungsdienste	178
9.2.5.3	Zuweisung von Deskriptoren	181
10	Ein Ansatz, der funktioniert	185
10.1	Computerhilfen bei hierarchisch geordneten Deskriptoren	185
10.2	Computerhilfen bei anders als hierarchisch geordneten Deskriptoren	201
10.3	Computerhilfen bei einfachen alphabetischen Konkordanzlisten	205
10.4	Computerhilfen bei bereichsabgrenzenden und relationsanzeigenden Deskriptoren (range and relational descriptors)	207
10.5	Rechnerunterstützte Recherche	208
	Nachwort zur Gegenwart	215
	Literaturquellen	219