

INHALTSVERZEICHNIS

INHALTSVERZEICHNIS.....	I
ABBILDUNGSVERZEICHNIS.....	V
TABELLENVERZEICHNIS.....	VII
BEZEICHNUNGEN	IX
DEFINITIONEN	XIII

KAPITEL 1

EINLEITUNG UND MOTIVATION	1
1.1 Übersicht	1
1.2 Aufbau der Arbeit	3
1.3 Motivation	5

KAPITEL 2

EINE GRUNDLEGENDE BETRACHTUNG DES INFORMATION RETRIEVAL	9
2.1 Überblick	9
2.2 Definition des Information Retrieval.....	9
2.3 Grundstruktur eines IR-Systems.....	11
2.4 Textanalyse	13
2.4.1 Zeichenkettenorientierter Ansatz	15
2.4.2 Statistischer Ansatz	17
2.4.3 Probabilistischer Ansatz.....	18
2.4.4 Linguistischer Ansatz	18
2.4.5 Wissensbasierter Ansatz.....	19
2.5 Bewertung von Retrievalsystemen	20
2.5.1 Bewertung der Systemeffizienz.....	21
2.5.2 Bewertung der Systemeffektivität	22
2.6 Modelle von Retrieval-Systemen	24
2.6.1 Exact-Match-Retrieval: Das Boolesche Retrieval	24

2.6.1.1	Boolesche Operatoren.....	26
2.6.2	Best-Match-Retrieval.....	27
2.7	Funktionen zur Verbesserung der Retrievalergebnisse.....	27
2.7.1	Clustering von Dokumenten.....	28
2.7.2	Relevanz Feedback.....	30
2.7.3	Visualisierung von Suchresultaten.	31
2.7.3.1	Beispiele für Visualisierungstechniken: Lyberworld	32
2.7.3.2	Beispiel zur Modellierungssprache: VRML.....	35
2.7.4	Metasprachen.....	36

KAPITEL 3

PROBLEMANALYSE DER SUCHE IM INTERNET.....		39
3.1	Übersicht	39
3.2	Das Internet	39
3.3	Problemstellung der Suche im Internet.....	41
3.4	Kriterien zur Beurteilung der Suchdienste	44

KAPITEL 4

DAS RECHERCHIEREN IM INTERNET		49
4.1	Überblick	49
4.2	Klassifikation der Recherchewerkzeuge im WWW.....	49
4.3	Einfache Search Engine.	52
4.4	Automatische Suchwerkzeuge im WWW: Architektur einer Search Engine	55
4.4.1	Roboter: Definition und Klassifizierung von Robotern.....	56
4.4.1.1	Vor- und Nachteile von Robotern	57
4.4.2	Datenbank mit einem indexierenden Werkzeug einer WWW- Search Engine.....	58
4.4.3	Benutzeroberfläche und die Abfragesprache einer WWW- Search Engine.....	59
4.5	Meta-Suchmaschine	60
4.5.1	Beispiel Meta-Service Suche: MetaCrawler.....	61
4.6	Individuelle Suche.....	62
4.6.1	Ein Beispiel für eine individuelle Suche: Fish-search.....	62
4.7	Verteilte Suche.....	64
4.7.1	Beispiel einer verteilten Suche: Ingrid.....	64

4.7.2	Beispiel einer verteilten Suche: Harvest	67
4.7.3	Beispiel einer verteilten Suche: Aliweb	70

KAPITEL 5

KONZEPTION DES DOKUMENTENMANAGEMENTS VON IVIRES		73
5.1	Überblick	73
5.2	IVIRES im Vergleich zu anderen Systemen im Internet.....	73
5.3	Grundelemente von IVIRES	76
5.4	Das Dokumentenmanagement	78
5.4.1	Das Aufnehmen eines Dokuments	79
5.4.2	Das Recherchieren	85

KAPITEL 6

SYSTEMENTWURF: DIE TEXTANALYSE		87
6.1	Übersicht	87
6.2	Inhaltliche Erschließung.....	88
6.2.1	Gewichtung nach Begriffshäufigkeiten.....	89
6.3	Suchoptionen	92
6.3.1	Vergabe des Metavektors	94
6.4	Clusteranalyse	95
6.4.1	Euklidische Distanz	95
6.4.1.1	Normierung der Dokumentvektoren	98
6.5	K-Means-Verfahren.....	100
6.6	Dateiorganisation	104

KAPITEL 7

SYSTEMENTWURF: DAS RETRIEVAL.....		107
7.1	Überblick	107
7.2	Das Konzept der Relevanz	107
7.2.1	Retrieval Status Value (RSV).....	109
7.2.1.1	Das Filterungsmaß	109
7.2.1.2	Beispiel	113
7.2.1.3	Gegenüberstellung von Filterungsmaß und Cosinusmaß.....	114

7.3	Der iterative Suchprozeß: Relevanz- Feedback Verfahren	117
7.3.1	Modifikation der Suchanfrage.....	119
	Beispiel	123
7.4	Eine einheitliche Integration der verteilten Indexe:	
	Weitere Ebene der Brokerhierarchie.....	124
7.5	Das Retrieval: Selektion der Datenbanken	128
 KAPITEL 8		
	ZUSAMMENFASSUNG UND ZUKÜNFTIGE ARBEIT	133
 LITERATURVERZEICHNIS.....		137
INTERNET-ADRESSEN.....		147
ADRESSE DER SUCHDIENSTE.....		149