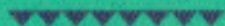




Automated Information Retrieval



Theory and Methods



Valery I. Frants
Jacob Shapiro
Vladimir G. Voiskunskii

Soranaree University of Technology



31051000465357



Contents

Preface

xi

1

The System and the Systems Approach

1.1	Introduction	1
1.2	The Notion of "System"	1
1.3	The Main Characteristics of the Systems Approach	4
1.4	Motion	7
1.5	Input and Output Quantities	9
1.6	Modeling	12
1.7	Control	15
1.8	System Design	22
1.9	Conclusion	25
	References	25
	Bibliographic Remarks	26

2

Vital Activities and Needs

2.1	Introduction	27
2.2	The Individual as a Functioning System	28
2.3	The Information Need	34
2.4	Information	40
2.5	Document	46
2.6	Conclusion	48
	References	49
	Bibliographic Remarks	50

3**Information Crisis**

3.1	Introduction	51
3.2	Information Activity and Information Crises	52
3.3	The Modern Information Crisis	56
3.4	Search for Information	61
3.5	Conclusion	66
	References	66
	Bibliographic Remarks	67

4**Concept of an Information Retrieval System**

4.1	Introduction	68
4.2	IR System as Object	68
4.3	The Purpose of Creating an IR System	70
4.4	Function of a Documentary IR System	74
4.5	Structure of a Documentary IR System	78
4.6	Conclusion	87
	References	88
	Bibliographic Remarks	88

5**Information Retrieval Language**

5.1	Introduction	89
5.2	Languages and Methods of Information Retrieval	89
5.3	IRL for the IR System	96
5.4	Some Tendencies of IRL Development	103
5.5	Automatic Methods of Descriptor Dictionary Compilation	107
5.6	Semantic and Grammatical Components of IRL	113
5.7	Relevance, Pertinence, Recall, and Precision	128
5.8	Conclusion	133
	References	134
	Bibliographic Remarks	135

6**Automatic Indexing of Documents**

6.1	Introduction	136
6.2	On the Problem of Indexing	137

6.3	Main Directions in Automatic Indexing	141
6.4	Some Methods of Text Analysis for Automatic Indexing Using a Dictionary	143
6.5	Algorithms of the Automatic Indexing of Documents	152
6.6	Statistical Indexing Methods	155
6.7	Some Issues in Automatic Indexing	158
6.8	Conclusion	163
	References	164
	Bibliographic Remarks	165

7

Automatic Indexing of Search Requests

7.1	Introduction	166
7.2	Some Aspects of Constructing Query Formulations	167
7.3	Approaches to the Automatic Indexing of Search Requests	173
7.4	Algorithm for Automatic Construction of Query Formulations in Boolean Form	185
7.5	The Choice of Bound Values	193
7.6	Some Aspects of Algorithm Functioning in Information Retrieval Systems	194
7.7	Conclusion	197
	References	198
	Bibliographic Remarks	199

8

Storage and Access to Information

8.1	Introduction	201
8.2	Sequential File and Sequential Access	202
8.3	Primary and Secondary Storage	204
8.4	Relative File and Random Access	210
8.5	Inverted Files	214
8.6	Conclusion	220
	References	221
	Bibliographic Remarks	221

9

Control and Feedback in IR Systems

9.1	Introduction	222
9.2	Some Questions for Constructing Adaptive IR Systems	223

9.3	Some Approaches to the Realization of Adaptive Feedback	229
9.4	Feedback Algorithm for the Static Collection of Documents	234
9.5	Feedback Algorithm for the Dynamic Collection of Documents	237
9.6	Evaluation	239
9.7	Feedback for Search Optimization	242
9.8	Criterion for Selection of the Best System's State	246
9.9	Selective Algorithm for the Static Collection of Documents	249
9.10	Selective Feedback Algorithm for Dynamic Collection	252
9.11	Internal Control	255
9.12	Conclusion	256
	References	258
	Bibliographic Remarks	259

10

Evaluation of Search Results

10.1	Introduction	260
10.2	Aspects of Evaluating	261
10.3	Problems of Evaluating the Functional Effectiveness of a Document Search	267
10.4	Limits of Applicability of Complex Search Characteristics	271
10.5	Determination of Recall	279
10.6	Construction of Complex Search Characteristics	281
10.7	"Physical Meaning" of Complex Search Characteristics	292
10.8	Order Preservation Property	293
10.9	Fuzzy Scales of Pertinence	300
10.10	Another Formulation of the Goal of Document Search	310
10.11	Conclusion	314
	References	315
	Bibliographic Remarks	316

11

Evaluation of Macroevaluated Objects

11.1	Introduction	318
11.2	Determination of Expediency for Evaluating Specific Macroevaluated Objects	319
11.3	Averaging Values of Functional Effectiveness	324
11.4	Requirements for Search Characteristics	325
11.5	Comparative Evaluation of Macroevaluated Objects	328
11.6	Some Experiments on the Evaluation of Macroevaluated Objects	329

11.7	Conclusion	333
	References	334

12

Some Directions in the Development of IR Systems

12.1	Introduction	335
12.2	IR Systems and Artificial Intelligence	335
12.3	Satisfying IN: Additional Possibilities	344
12.4	The Effects of IN Components on the IR System	348
12.5	Future Developments in Considering IN Components	354
12.6	Conclusion	355
	References	356
	Bibliographic Remarks	357

Index	359
-------	-----