British Library Research and Development Report No. 5547

University of Aston, Computer Centre
Technical Report No. TR80002

DESIGN STUDY FOR AN ANOMALOUS STATE OF KNOWLEDGE BASED INFORMATION RETRIEVAL SYSTEM

Final Report on Grant S1/SG/09
2 May 1978 - 15 June 1979

N.J. Belkin
Centre for Information Science
The City University

R.N. Oddy
Computer Centre
University of Aston in Birmingham

September 1979

All opinions and observations are those of the authors, and not necessarily those of the British Library.

TABLE OF CONTENTS

		Page
ABSTRACT		1
Chapter 1.	INTRODUCTION	
1.1	Project origins	5
1.2	Background	6
1.3	Project specification	7
1.4	Aims of the Design Study	8
Chapter 2.	THEORETICAL BASIS OF THE ASK IR SYSTEM	11
Chapter 3.	METHODS	
3.1	General experimental outline	16
3.2	Problem statements	16
3.3	Abstracts	17
3.4	Evaluation procedures	18
3.5	Text analysis - principles	18
3.6	Text analysis - procedure	21
Chapter 4.	RESULTS	
4.1	Problem statements - general characteristics	31
4.2	Problem statements - evaluation	35
4.3	Abstracts - general characteristics	40
4.4	Abstracts - evaluation	40
Chapter 5.	DATA ANALYSIS	
5.1	Classification of ASKs	44
5.2	Retrieval strategies	45
Chapter 6.	DISCUSSION	
6.1	Design problems	52
6.2	Text characteristics	53
6.3	Evaluation of representations	53
6.4	Text analysis	54

		Page
Chapter 7.	CONCLUSIONS AND FURTHER RESEARCH	56
	ACKNOWLEDGEMENTS	57
	REFERENCES	58
Appendix A	Evaluation package for problem statements	60
Appendix B	Evaluation package for abstracts	65
Appendix C	Summaries of problem statements	68
Appendix D	Association map representations of problem statements	87
Appendix E	Single-link clustering representation of problem statements	123
Appendix F	Abstracts	159
Appendix G	Association map representations of abstracts	195

FIGURES

1.	An ASK-based IR system	9
2.	A model of IR	12
3.	The transcript of a problem statement	20
4.	Token processing	23
5.	Calculation of association strengths	24
6.	The strongest associations from the problem statement in Figure 3	25
7.	Association map for problem statement in Figure 3	27
8.	Association clusters for problem statement in Figure 3	28
9.	An example abstract	29
10.	Association map for abstract in Figure 9	30
11.	Condensed problem statement network - 1	48
12.	Condensed problem statement network - 2	49

TABLES

1.	Token - token and type - token ratios for oral problem statements.	32
2.	Association strengths and number of types for oral problem statements	33
3.	Token-token and type-token ratios for written problem statements	34

		Page
4.	Association strengths and number of types for written problem statements	34
5.	Subject areas of interviewees	36
6.	Association map evaluation (Problem statement)	37
7.	Association cluster evaluation (Problem statement)	38
8.	Format comparison	39
9.	Token-token and type-token ratios for abstracts	41
10.	Association strengths and number of types for abstracts	42
11.	Abstract representation evaluation	43
12.	Summary of ASK types	46

ABBREVIATIONS

ASK	Anomalous state of knowledge
CIS	Central Information Services, University of London
IR	Information retrieval
PS	Problem structure