

Okapi at City

An evaluation facility for interactive IR

Stephen Walker and Micheline Hancock-Beaulieu

with contributions by

Ayşe Göker
Lee McCluskey
Scarlett Palmer

Centre for Interactive Systems Research
City University
London
August 1991

British Library Research Report 6056

© British Library Board 1992

Contents

1	Introduction	6
1.1	Introduction	6
1.2	Background	6
1.3	Development of Okapi at City University	7
1.3.1	The PCL Okapi system	7
1.3.2	Parameterization	8
1.3.3	Updating procedures	9
1.3.4	Databases	9
1.3.5	Logging	10
1.4	Installation and use of Okapi at City	10
1.4.1	Installation	10
1.4.2	Use	11
2	Query expansion in a library catalogue	13
2.1	Introduction	13
2.1.1	Automatic query expansion	13
2.1.2	Installation of catalogue system	14
2.2	Background	14
2.2.1	Library database	14
2.2.2	Library setting	15
2.2.3	Network access	15
2.3	Evaluation methodology	15
2.3.1	Transaction logs	15
2.3.2	Search replays	16
2.3.3	Pre-search and post-search questionnaires	16
2.3.4	Post-search interviews	17
2.3.5	Structured interviews of network catalogue users	17
2.4	Findings	18
2.4.1	Overall usage of query expansion facility	18
2.4.2	Effectiveness of query expansion	18
2.4.3	Perceived usefulness of query expansion facility	19
2.4.4	Search intentions	20
2.4.5	Replay of searches with “MORE” option not used	20
2.4.6	Previous experience with online catalogues	21
2.4.7	Ease of use	21
2.4.8	Help facility	22

2.4.9	User satisfaction with search outcome	22
2.4.10	Browsing references	23
3	Transaction log analysis	24
3.1	Introduction	24
3.1.1	Source of the data	24
3.1.2	Sessions and searches	25
3.1.3	Transaction logs	25
3.2	Sessions and searches	27
3.2.1	Aborted searches	27
3.2.2	Failed searches	28
3.3	Records retrieved, displayed and chosen	28
3.3.1	Numbers of records retrieved	28
3.3.2	Effect of number of records reported on record display	29
3.3.3	Full records	30
3.3.4	Relevance judgements	31
3.4	Query expansion	32
3.4.1	Takeup of query expansion option	32
3.4.2	Results of query expansion searches	33
3.4.3	Source of records displayed and chosen	34
3.4.4	How useful was query expansion?	34
3.5	System use patterns	36
3.6	Differences between classes of users	36
3.7	Best match searching	37
4	Frequent users	39
4.1	Introduction	39
4.2	Obtaining the data	39
4.2.1	Background information about users	40
4.2.2	Individuals' use of language	41
4.2.3	Search language in general	41
4.3	Analysing the data	41
4.4	Results	42
4.4.1	General points about frequent users	42
4.4.2	Suggestions for system modifications	42
5	Towards an adaptive IR system	44
5.1	Introduction	44
5.2	The probabilistic model for an IRS	45
5.3	The learning component	46
5.3.1	Inputs to the learner	46
5.3.2	Formulating and evolving the context	47
5.4	Testing context formulation	49
5.4.1	The Okapi system	49
5.4.2	An example	49
5.5	Conclusions and future work	51

<i>CONTENTS</i>	3
6 Conclusions	52
6.1 Automatic query expansion	52
6.1.1 Query expansion in the library system	53
6.1.2 Query expansion in searching INSPEC	54
6.1.3 Heuristics for query expansion	55
6.1.4 Query expansion: conclusions	55
6.2 Towards adaptive IR systems	57
A Okapi system description	59
B Okapi transaction logs	71
C Questionnaires	77
C.1 Questionnaires for library users	77
C.1.1 Exhaustiveness of search	77
C.1.2 Previous experience	77
C.1.3 Post-search questionnaire	78
C.1.4 Structured interview	78
C.2 Questionnaire for network users	79
D Notes on a frequent user	80

List of Figures

3.1 Transaction log summary	26
A.1 Welcome screen	59
A.2 Information window for welcome screen	60
A.3 Information offered after user identification	60
A.4 Search input screen	61
A.5 Information offered from search input screen	61
A.6 Search results screen	61
A.7 Replacement of misspelt word	62
A.8 Brief record display	62
A.9 Options from brief display	63
A.10 Full record display	63
A.11 Brief record display with “MORE” option	64
A.12 Results of query expansion	65
A.13 Brief display from query expansion search	65
A.14 Record from query expansion	66
A.15	67
A.16 Records from second iteration of query expansion	67
A.17	68
A.18 Options after query expansion	68
A.19 Choosing the PRINT option	69
A.20 Editing a search	69
A.21 Full record display: INSPEC	70
A.22 Brief record display: INSPEC	70

List of Tables

2.1	Availability and usage of query expansion	18
2.2	Usage of the “MORE” option and search sessions	18
2.3	Items selected with and without query expansion	19
2.4	Items displayed and selected before and after query expansion	19
2.5	Items selected per search after query expansion	19
2.6	Search intentions	20
2.7	Query expansion in replayed searches	21
2.8	Users’ experience of online catalogues	21
2.9	Ease of use of Okapi	21
2.10	Ease of use of Okapi compared to CLSI	22
2.11	User satisfaction	22
2.12	Perceived number of references looked at	23
3.1	Sessions and searches	27
3.2	Distribution of number of records retrieved	28
3.3	Percent of non-failed searches in which no records were displayed	29
3.4	Action following search with no records displayed (CAT1 set)	30
3.5	Full records displayed and chosen: searches in which some brief records were displayed	31
3.6	Full records displayed and chosen: searches in which some full records were displayed	31
3.7	Takeup of query expansion by search result	33
3.8	Takeup of query expansion by number of records chosen . . .	33
3.9	Records displayed and chosen from query expansion	34
3.10	Sources of displayed and chosen records: CAT1	35
3.11	Sources of displayed and chosen records: INSPEC	36
3.12	System use Jan–Apr 1991: all identified users	37
3.13	System use Jan–Apr 1991: people who had used system be- fore Jan 1 1991	37
5.1	50