

Contents

Foreword Tony Byrne **xi**

Preface **xiii**

Acknowledgements **xvii**

Making search work - critical success factors **xix**

1 Search must work 1

Searching but not finding 1
The advent of Eureka 5
The four dimensions of search 6
Findability 9
Search and navigation 10
The role of the Information Discovery Manager 11
References 12

2 How search works 13

Introduction 13
Content definition 14
Indexing 15
Taxonomy management 17
Stemming and lemmatization 18
Document confidentiality 19
Query management 21
Ranking and relevance 22
Vector space model 23
Meaning-based computing 24
Manipulating relevance algorithms 25
Results formatting 25
Document access 26
Comparative evaluation of search software 27

Text mining 28
References 28

3 The search business 31

The market for search 31
Search appliances 33
Hosted search 33
Google 34
IBM 36
Microsoft 37
Oracle 38
SAP 38
Other search software vendors 38
References 39

4 Making a business case for search 41

Introduction 41
Total cost of implementation 42
Price-performance 46
Making a business case 47

5 Specifying and selecting a search engine 49

Introduction 49
Using personas to specify user requirements 50
Determine initial user segmentation 51
Document audit 52
Concurrent users 55
Document security 55
IT architecture considerations 56
Scalability 57
The need for a team effort 58
Schedule for the selection process 59
Knowing what you don't know 59
Developing the request for a proposal 61
Information from the vendor 62
Open source software 63
Pilot testing 64
References 64

6 Optimizing search performance 65

Introduction 65
Search system performance 66

System availability	67
Search engine tuning	68
Privacy issues	71
References	73

7 Search usability 75

Supporting information discovery	75
The query box	76
Search hit presentation	77
Search page presentation	80
Results presentation	81
Some examples of good search presentation	83
Accessibility issues	87
References	87

8 Desktop search 89

Introduction	89
The role of desktop search	90
Selecting a desktop search product	91
References	94

9 Implementing web search 95

Introduction	95
The ten-step process to implementing web search	96
Reference	103

10 Implementing search for an intranet 105

The ten-step process to implementing intranet search	105
References	111

11 Enterprise search 113

Introduction	113
Document diversity	114
Structured database integration	115
Information technology requirements	116
Managing the implementation	116
Managing the search team	117
Managing the search desktop	118
The budget challenge	118
Federated search	119
Ten steps to implementing enterprise search	120

12 Multilingual search 127

- Searching the Tower of Babel 127
- Searching multiple languages 128
- Unicode 129
- Identifying index terms 129
- Transliteration of names 130
- Understanding the language search options 132
- References 133

13 Future directions 135

- Search is going to be a big business 136
- Search software vendors will continue to show low profitability 136
- The web will continue to stimulate search innovation 137
- CMS and search will become better integrated 138
- Search solutions that work across semistructured and structured data will be much in demand 138
- Business intelligence and search will converge 139
- Federated search will continue to improve, because it has to 139
- Visualization software will improve rapidly 140
- There will be substantial progress in searching single and sequential images 140
- Social tagging will come to search 140
- Third-party information vendors will have to review their position on search 140
- Reference 141

Appendix: Search software vendors 143**Further reading 159****Glossary 165****Subject index 167****Company index 171**