

Contents

Foreword

Acknowledgements

Preface

Part I: Setting the scene

1. Ensuring Success, Reducing Risk³

Introduction

In the beginning

Exponential growth

The rise of enterprise search

The start of the journey

2. Why Enterprise Search is Difficult

Expectations

What is enterprise search?

What makes search a challenge?

Management challenges

Technical challenges

Scaling

Upgrading

Tuning

Customising

Transforming

Informational issues

Versions

Exceptions

Volatility

Cracking the tough nut

3. The Enterprise Search Industry

Market overview

Google and the enterprise

Pressure and compression

Business intelligence⁶

Superplatforms

Competitive implications

eDiscovery and brand names

eDiscovery

Brand name vendors

Fast-growing sectors

Specialist systems

Up-and-coming vendors

Challenges for eDiscovery and brand name vendors

New players in enterprise search

Low-cost and open source systems

Search utilities

Start-ups

Characteristics of the enterprise search market

Implications

Part II: How search works

4. The Technology of Search

Introduction

Components of a search system

Crawler sub-system

Document processing sub-system

Indexing sub-system

Query processing sub-system

Administrative sub-system

Infrastructure

Technology to handle peaks and valleys

Assisted navigation

Adding advanced text processes

Content processing

Semantic processing

Entity extraction

Classification

5. Emerging Search Technologies

New technologies

Discovery

Federated search

Social search

Semantic search

What about ontologies, taxonomies and classification systems?

Classification system

Knowledge base

Ontology

Taxonomy

Keeping perspective

Practical suggestions regarding search technology

6. Managing Security

Collection and document level security

Organisation-wide policy

Access control

Approaches to secure search

Vendor approaches to security

Managing data privacy

Secure search administration

Programming policies for security

A search security checklist

Define the system scope
Plan, modify, follow the plan, adapt, and then repeat
Players and constituencies
Security metrics
Specify thresholds and parameters
Perform security audits

7. Managing Language

Introduction

The challenges posed by the real world of language
Handling multiple languages
Same words – different meanings
Specific challenges for search technology

Linguistic analysis

Latent semantic indexing (LSI)

Entity extraction

Searching for personal names

De-duplication

Managing collections in multiple languages

Part III: How to select a vendor

8. Defining Your Enterprise Search Requirements

Introduction

Total cost of search implementation

License fee

Software maintenance

Upgrades

Additional modules

Database licenses

Professional services

Hardware and infrastructure

Training

In-house support team

Making a business case

Defining user requirements

Current search system audit

Enterprise application road map

Search screens

Collection analysis

Federated search requirements

Repository method

9. Writing the Tender Specification

Introduction

Defining objectives

Information systems strategy

Compiling the RFP

Accessibility support

APIs

Crawls and indexing

File types

Proof of concept

Risk assessment

Security management

Upgrade release schedule

Supporting a global implementation

Using the incumbent systems integrator

Using a new systems integrator

Using a specialist search integrator

10. Selecting a Vendor

Introduction

The selection team

Developing a short list

The presentation

The difficult questions

Additional software licenses

APIs

Customisation

Hardware and software

License terms

Preferred architecture

Search reporting

Security

Taxonomy management

Training and support

The debrief

11. Proof of Concept

Methods exist, not followed

A definition

Steps in the proof of concept

Overall work flow

Management track

Test corpus track

Infrastructure track

Case example: demos are not systems

A plan

Test corpus

Document distribution and collection

Special cases

Legacy data

Confidential data

Copyrighted information

Practical considerations

A General Mills bake off

What to test

Test queries

Avoiding pitfalls

Reducing risk

Part IV: How to implement search

12. Managing the Search Team

Introduction

At the very beginning

Roles and responsibilities

Search (Information Discovery) Manager

Search Technology Manager

Information Specialist

Search Analytics Manager

Search Support Manager

The core team

Supporting global enterprise search

Creating a global Center of Search Excellence

Reporting lines

13. Deploying the Search Solution

Scope

What to do

Why search is different

Information is fluid

Search is not one system

Content changes

Dependencies

Two dependency hot spots

The diagram

The coding loop

The security loop

Avoiding delays and cost overruns

Lack of management support

Too much customisation

Content woes

Tips and tricks

Plan early and adapt to changes

Obtain management support

Limit features

Minimise customization

Figure out the content scope

Security

Use consultants

14. Search Usability

Introduction

Ranking and relevance

Other search metrics

The search dialogue

Federated search

- Search page design
 - Limiting the scope of the search*
 - Numeric values*
 - Other common search problems*
 - Best bets and URL management*
 - Auto categorization*
 - User-centric design*

15. Search Analytics

- Introduction
- A multiplicity of parameters
- Search logs
 - Number of searches per hour/day*
 - Top 50 searches by search terms/query*
 - Top 50 searches leading to no results being presented*
 - Top 50 searches leading to no document being selected*
 - Top 50 most requested documents*
 - Top 50 searches where more than three pages of hits were presented*
 - Top 50 searches by duration of search*
 - Searches by number of results presented*
 - Saved searches*
 - Searches where new documents are alerted to the user*
- It takes more than search logs
- Taking action

16. Search System Upgrades

- Backstory
- Case example: major vendor
- Case example: Microsoft SharePoint
- The dilemma
- Trade-off analysis needed
 - Data you need*
 - Past history*
 - What other licensees tell you*
 - Vendor's sales or support engineers*
 - The web*
 - Third-party search experts*
 - Interpreting the data*
- Reasons to upgrade
- Reasons not to upgrade
- A road map
 - Spell out what's to be done*
 - Plan for a restore or fail over*
 - Get commitment from your team*
 - Cut over day*
 - Understanding the storage method*
 - Key points to note*
- Net net

Part V: Wrap-up

17. Action This Day

Introduction

A growing appetite

A structured approach

Planning

Preparation

Implementation

Adaptation

A strategic perspective

**Appendix 1 – Summary of Key Questions for Successful
Enterprise Search**

Glossary

The Authors