

Contents

Introduction	xxvii
Chapter 1: Introduction to Enterprise Search	1
Why Enterprise Search	1
A Tale of Two Content Types	1
Security, Security, Security	2
Algorithms to the Rescue	2
We All Love the Web and HTTP	3
Conclusion	4
Chapter 2: Overview of Microsoft Enterprise Search Products	5
Enterprise Search Product Overviews	5
Windows Desktop Search/Windows Vista	5
Features in Windows Vista Search	6
Windows SharePoint Services	11
SharePoint Search Architecture	11
Crawling Content	12
Searching Content	13
Configuring Search	14
Platform Services	14
Microsoft Search Server 2008	16
Simplified Setup and Administration	16
Federation Capabilities	18
Different Editions of Search Server 2008	19
What about WSS and Microsoft Office SharePoint Server?	20
Microsoft Office SharePoint Server	20
People Search	20
Business Data Catalog	21
The Microsoft Filter Pack	23
Connectors for Documentum and FileNet	23
Windows Live	24
FAST and SharePoint	25
Other Server Products (Exchange, SQL)	25
Conclusion	25

Chapter 3: Planning and Deploying an Enterprise Search Solution	27
Key Components	27
The Index Role	27
The Query Role	28
The Shared Services Provider	28
The Database Server	29
Search Topologies	29
Single Server	29
A Small Farm	30
A Three-Server Farm	31
A Medium Server Farm	31
Larger Farms	32
Search Software Boundaries	33
Hardware Sizing Considerations	34
The Index Server	35
Query Servers	37
Database Servers	37
Testing	38
Performance Monitoring	39
Search Backups	41
Index Server Recovery Options	42
Using Federation to scale?	43
Conclusion	44
Chapter 4: Configuring and Administering Search	45
Configuring Search from Central Administration	45
The Search Services	45
The Office SharePoint Server Search Service	46
Windows SharePoint Services Search	50
Manage Search Service	52
Manage Content Database — Search Server 2008	57
Configuring Search from the Shared Services Provider	58
Creating or Editing the SSP Settings	59
SSP Search Administration	59
The Default Content Source	60
Full versus Incremental Crawls	62
Search Schedule	63
Additional Content Sources	64
Interacting with a Content Source	65
Crawl Rules	66

Crawl Logs	70
File Types	72
Reset All Crawled Content	72
Search Alerts	72
Authoritative Pages	72
Federated Locations	73
Managed Properties	73
Shared Search Scopes	75
Server Name Mappings	77
Search Result Removal	78
Search Reporting	78
The Other Search Settings	79
Configuring Search Locally on the Server	80
IFilters	80
Installing the Microsoft Filter Pack	80
Maximum Crawl Size	82
Reset the Search Services	82
Crawling Case-Sensitive Web Sites	82
Diacritic-Sensitive Search	82
Conclusion	83
Chapter 5: Searching LOB Systems with the BDC	85
BDC Architecture and Benefits	85
The Application Definition File	87
XSD Schema File	87
BDC Definition Editor Tool	87
BDC Metadata Model Overview	88
MetadataObject Base Class	89
LobSystem	89
LobSystemInstances and LobSystemInstance	92
Entities and Entity Element	94
Identifiers and Identifier Element	95
Methods and Method Element	96
Parameters and Parameter Element	98
FilterDescriptors and FilterDescriptor Element	98
Actions, Action, and ActionParameter Elements	101
MethodInstance Element	101
TypeDescriptors, TypeDescriptor, DefaultValue Elements	106
Associations and Association Element	108
Complete BDC XML Samples	109

Contents

BDC Web Parts, Lists, and Filters	109
Business Data List Web Part	110
Business Data Related List Web Part	110
Business Data Item Web Part	111
Business Data Actions Web Part	112
Business Data Item Builder Web Part	112
BDC in SharePoint Lists	113
Modifying Your BDC Profile Page	116
Searching the BDC	117
Adding a Content Source for Crawling BDC Data	117
Mapping Crawled Properties	119
Create a Search Scope	119
SharePoint Designer and the BDC	120
The BDC API	122
The BDC Assemblies	122
The Microsoft.Office.Server Namespaces	123
Putting It Together: Building Custom Applications for the BDC	123
Connecting to the Shared Services Database	125
Displaying LOBSystemInstances	125
Working with Entities	125
Working with an Entity – Finders, Fields, and Methods	126
Executing a Method and Displaying the Results	126
Working with Associations and Actions	128
TroubleShooting the BDC	129
Conclusion	129
Chapter 6: User Profiles and People Search	131
User Profiles	131
Managing User Profiles	133
Profile Services Connections	133
Configuring Profile Imports	134
Profile Properties	138
Configuring Profile Properties	139
BDC Supplemental Properties	142
Configuring for BDC Import	143
People Search	145
The Search Center	145
People Search Page and Tab	146
Results Page	147

Customizing People Search	150
Customizing the Advanced Search Options Pane	151
Adding Custom Properties to the People Results Page	152
Summary	154

Chapter 7: Extending Search with Federation **155**

The Concept	156
Crawl and Index or Federate?	158
Unified Result Set	158
Freshness of Results	158
Avoiding “Double Indexing” Content	159
Content outside Your Firewall	159
Federated Search Locations	160
Triggers	163
Always	164
Prefix	164
Pattern	164
Using Named Capture Groups to Send Only Specific Query Terms	165
Location Type	168
Query Template	168
OpenSearch 1.0/1.1	168
Required and Optional Query Template Tokens	170
Search Index on this Server	171
More Results Link	172
Display Information	172
Specifying a Branding Icon	174
Customizing the Title of a Federated Results Web Part	174
Writing Custom XSL	175
Properties	175
Sample Data	176
Restrictions	177
Credentials	177
Anonymous	178
Common	178
User	178
Configuring Search Server 2008 for “User” Authentication	179
Configuring SharePoint Server for User Authentication	180
Security Trimming Federated Results	180

Contents

Federation Web Parts	182
Federated Search Results Web Part	183
Federated Search Results Web Part Properties	185
Location	185
Results per Page	186
Use Location Visualization	186
Retrieve Results Asynchronously	186
Results Query Options	186
Show More Results Link	187
Appearance Settings	187
Caching Location Settings	187
Top Federated Results Web Part	187
Multiple Locations	187
Retrieve Results Asynchronously	188
Summary	188
Chapter 8: Securing Your Search Results	189
Security Architecture in SharePoint	189
Best Bets	191
Controlling Indexing to Secure Content	191
What about IRM-Protected Documents?	191
Custom Security Trimmers	191
Implementing a Custom Security Trimmer	192
Registering the Custom Security Trimmer	193
Performance Considerations	193
Bringing It All Together: Building a Custom Security Trimmer	194
Getting Started	194
Signing Your DLL and Adding It to the GAC	198
Registering your Customer Security Trimmer	198
Debugging Your Customer Security Trimmer	200
BDC Security Trimming	201
Authentication and the BDC	201
BDC and Search	202
Performance Implications	202
Writing Your Custom Security Trimmer	202
Deploying Your BDC Security Trimmer	206
Debugging Your BDC Security Trimmer	209
Securing Your Search Server 2008	212
Default Content Access Account	212
Single Server Deployment	212
Server Farm Deployment	212
Conclusion	213

Chapter 9: Customizing the Search Experience	215
Describing the Flow of a Typical Search	215
Customizing the Search Center — No Code	218
Tab User Interface	219
Thesaurus and Synonyms	222
Restarting the SharePoint Search Service	224
Stemming	224
Customizing People Results	225
SharePoint Designer Support	225
Customizing your Search Results — XSLT	225
Stepping through the XSLT	227
Root Template	227
Parameter Section	228
No Search Results	229
Search Results	230
Hit Highlighting Template	230
Display Size Template	230
Display String Template	231
Document Collapsing Template	231
More Results for Fixed Query	231
Discovered Definitions	231
Working with SPD to Create Your XSLT	231
Customizing Hit Highlighting in Search	235
Adding and Rendering Custom Properties	237
Adding a New Managed Property	241
Customizing Core and Advanced Search Results	243
Customizing People Search and Results with Custom Properties	246
Using Fixed Queries	251
Adding Custom Web Parts	251
Search Community Toolkit	252
Conclusion	252
Chapter 10: Understanding and Tuning Relevance	253
What Is Relevance?	253
Built-In Ranking Elements and Algorithms	254
Understanding and Tuning Relevance	254
Things You Can Change outside of the Ranking Algorithms	255
Content Management	255
Language	260
Understanding Users' Query Behavior with Query Logs	269

Contents

User Interface	270
High-Confidence Matches	273
Things You Can Change That Directly Affect the Ranking Algorithm	280
Authoritative and Nonauthoritative Pages	280
Property Weighting	282
Global Ranking Parameters	284
Conclusion	287
Chapter 11: Building Applications with the Search API and Web Services	289
<hr/>	
Getting Started with References	289
Administration Object Model	289
SearchContext Object	290
Content Sources	291
Working with the Content Class	291
Creating a New Content Source	293
Custom, File Share, Exchange, and Lotus Notes	293
Business Data Catalog Content Source	293
Sample Code: Creating a Content Source	294
Enumerating Existing Content Sources	295
Deleting a Content Source	295
Crawling Content Sources	295
Working with Federated Locations	297
Crawl Rules and File Types	300
Creating a Crawl Rule	300
Working with File Types	302
Scopes	302
Schema Management	304
Querying All Categories	305
Querying All Crawled Properties	305
Querying All Managed Properties	306
Creating and Mapping a New Managed Property	306
Changing Property Weights	307
Editorial Results and Ranking	308
Working with Authoritative Sites and Rankings	310
Search Logs	311
Propagation	312
Writing Search Applications	312
Building Queries	312
Keyword Syntax for Searches	313
URL Syntax	314

SQL Syntax	315
Some Example Queries	318
The Query Object Models	320
.NET Query Object Model	320
Web Service Query Object Model	324
Gotchas When Using the Object Models	330
Tools to Help You Develop Search Applications	330
Content Query Web Part and Cross-List Queries	331
Conclusion	331
Index	333
