



Cognitive Computing and Big Data Analytics

Judith Hurwitz
Marcia Kaufman
Adrian Bowles

WILEY



Contents

Introduction	xvii
Chapter 1 The Foundation of Cognitive Computing	1
Cognitive Computing as a New Generation	2
The Uses of Cognitive Systems	2
What Makes a System Cognitive?	3
Gaining Insights from Data	4
Domains Where Cognitive Computing Is Well Suited	5
Artificial Intelligence as the Foundation of Cognitive Computing	6
Understanding Cognition	11
Two Systems of Judgment and Choice	12
System 1—Automatic Thinking: Intuition and Biases	13
System 2—Controlled, Rule-Centric, and Concentrated Effort	14
Understanding Complex Relationships Between Systems	15
Types of Adaptive Systems	16
The Elements of a Cognitive System	17
Infrastructure and Deployment Modalities	17
Data Access, Metadata, and Management Services	18
The Corpus, Taxonomies, and Data Catalogs	18
Data Analytics Services	18
Continuous Machine Learning	19
Hypothesis Generation and Evaluation	19
The Learning Process	19
Presentation and Visualization Services	20
Cognitive Applications	20
Summary	20

Chapter 2	Design Principles for Cognitive Systems	21
	Components of a Cognitive System	22
	Building the Corpus	23
	Corpus Management Regulatory and Security Considerations	25
	Bringing Data into the Cognitive System	26
	Leveraging Internal and External Data Sources	26
	Data Access and Feature Extraction Services	27
	Analytics Services	28
	Machine Learning	29
	Finding Patterns in Data	29
	Supervised Learning	29
	Reinforcement Learning	31
	Unsupervised Learning	32
	Hypotheses Generation and Scoring	33
	Hypothesis Generation	34
	Hypothesis Scoring	35
	Presentation and Visualization Services	36
	Infrastructure	37
	Summary	37
Chapter 3	Natural Language Processing in Support of a Cognitive System	39
	The Role of NLP in a Cognitive System	40
	The Importance of Context	40
	Connecting Words for Meaning	42
	Understanding Linguistics	43
	Language Identification and Tokenization	43
	Phonology	44
	Morphology	44
	Lexical Analysis	45
	Syntax and Syntactic Analysis	45
	Construction Grammars	46
	Discourse Analysis	46
	Pragmatics	47
	Techniques for Resolving Structural Ambiguity	47
	Importance of Hidden Markov Models	48
	Word-Sense Disambiguation (WSD)	49
	Semantic Web	50
	Applying Natural Language Technologies to Business Problems	50
	Enhancing the Shopping Experience	50
	Leveraging the Connected World of Internet of Things	51
	Voice of the Customer	51
	Fraud Detection	53
	Summary	53

Chapter 4	The Relationship Between Big Data and Cognitive Computing	55
	Dealing with Human-Generated Data	55
	Defining Big Data	56
	Volume, Variety, Velocity, and Veracity	56
	The Architectural Foundation for Big Data	57
	The Physical Foundation for Big Data	58
	Security Infrastructure	58
	Operational Databases	58
	Role of Structured and Unstructured Data	59
	Data Services and Tools	61
	Analytical Data Warehouses	61
	Big Data Analytics	62
	Hadoop	64
	Data in Motion and Streaming Data	67
	Analyzing Dark Data	68
	Integration of Big Data with Traditional Data	69
	Summary	70
Chapter 5	Representing Knowledge in Taxonomies and Ontologies	71
	Representing Knowledge	71
	Developing a Cognitive System	72
	Defining Taxonomies and Ontologies	73
	Explaining How to Represent Knowledge	75
	Managing Multiple Views of Knowledge	79
	Models for Knowledge Representation	80
	Taxonomies	80
	Ontologies	81
	Other Methods of Knowledge Representation	83
	Simple Trees	83
	The Semantic Web	83
	The Importance of Persistence and State	84
	Implementation Considerations	85
	Summary	85
Chapter 6	Applying Advanced Analytics to Cognitive Computing	87
	Advanced Analytics Is on a Path to Cognitive Computing	87
	Key Capabilities in Advanced Analytics	91
	The Relationship Between Statistics, Data Mining, and Machine Learning	92
	Using Machine Learning in the Analytics Process	93
	Supervised Learning	94
	Unsupervised Learning	96
	Predictive Analytics	98
	Business Value of Predictive Analytics	98
	Text Analytics	99
	Business Value of Text Analytics	100

	Image Analytics	101
	Speech Analytics	103
	Using Advanced Analytics to Create Value	104
	Building Value with In-memory Capabilities	105
	Impact of Open Source Tools on Advanced Analytics	106
	Summary	106
Chapter 7	The Role of Cloud and Distributed Computing in Cognitive Computing	109
	Leveraging Distributed Computing for Shared Resources	109
	Why Cloud Services Are Fundamental to Cognitive Computing Systems	110
	Characteristics of Cloud Computing	111
	Elasticity and Self-service Provisioning	111
	Scaling	111
	Distributed Processing	111
	Cloud Computing Models	112
	The Public Cloud	112
	The Private Cloud	114
	Managed Service Providers	114
	The Hybrid Cloud Model	115
	Delivery Models of the Cloud	117
	Infrastructure as a Service	117
	Virtualization	117
	Software-defined Environment	118
	Containers	118
	Software as a Service	118
	Platform as a Service	120
	Managing Workloads	120
	Security and Governance	121
	Data Integration and Management in the Cloud	122
	Summary	122
Chapter 8	The Business Implications of Cognitive Computing	125
	Preparing for Change	125
	Advantages of New Disruptive Models	126
	What Does Knowledge Mean to the Business?	127
	The Difference with a Cognitive Systems Approach	128
	Meshing Data Together Differently	129
	Using Business Knowledge to Plan for the Future	131
	Answering Business Questions in New Ways	134
	Building Business Specific Solutions	134
	Making Cognitive Computing a Reality	135
	How a Cognitive Application Can Change a Market	136
	Summary	136

Chapter 9	IBM’s Watson as a Cognitive System	137
	Watson Defined	137
	How Watson Is Different from Other Search Engines	138
	Advancing Research with a “Grand Challenge”	139
	Preparing Watson for <i>Jeopardy!</i>	140
	Preparing Watson for Commercial Applications	141
	Watson’s Software Architecture	142
	The Components of DeepQA Architecture	144
	Building the Watson Corpus: Answer and Evidence Sources	145
	Source Acquisition	146
	Source Transformation	146
	Source Expansion and Updates	147
	Question Analysis	148
	Slot Grammar Parser and Components for	
	Semantic Analysis	149
	Question Classification	150
	Hypothesis Generation	152
	Scoring and Confidence Estimation	153
	Summary	154
Chapter 10	The Process of Building a Cognitive Application	157
	The Emerging Cognitive Platform	158
	Defining the Objective	159
	Defining the Domain	160
	Understanding the Intended Users and Defining	
	their Attributes	161
	Defining Questions and Exploring Insights	162
	Typical Question-Answer Pairs	163
	Anticipatory Analytics	164
	Acquiring the Relevant Data Sources	166
	The Importance of Leveraging Structured Data Sources	166
	Analyzing Dark Data	167
	Leveraging External Data	167
	Creating and Refining the Corpora	168
	Preparing the Data	169
	Ingesting the Data	169
	Refining and Expanding the Corpora	170
	Governance of Data	171
	Training and Testing	171
	Summary	173
Chapter 11	Building a Cognitive Healthcare Application	175
	Foundations of Cognitive Computing for Healthcare	176
	Constituents in the Healthcare Ecosystem	177
	Learning from Patterns in Healthcare Data	179
	Building on a Foundation of Big Data Analytics	180
	Cognitive Applications across the Healthcare Ecosystem	181

Two Different Approaches to Emerging Cognitive Healthcare Applications	181
The Role of Healthcare Ontologies in a Cognitive Application	182
Starting with a Cognitive Application for Healthcare	183
Define the Questions Users will Ask	183
Ingest Content to Create the Corpus	184
Training the Cognitive System	185
Question Enrichment and Adding to the Corpus	185
Using Cognitive Applications to Improve Health and Wellness	186
Welltok	187
Overview of Welltok's Solution	187
CaféWell Concierge in Action	190
GenieMD	191
Consumer Health Data Platforms	191
Using a Cognitive Application to Enhance the Electronic Medical Record	191
Using a Cognitive Application to Improve Clinical Teaching	193
Summary	195
Chapter 12 Smarter Cities: Cognitive Computing in Government	197
How Cities Have Operated	197
The Characteristics of a Smart City	199
Collecting Data for Planning	200
Managing Operations	201
Managing Security and Threats	202
Managing Citizen-produced Documentation and Data	202
Data Integration Across Government Departments	203
The Rise of the Open Data Movement Will Fuel Cognitive Cities	204
The Internet of Everything and Smarter Cities	204
Understanding the Ownership and Value of Data	205
Cities Are Adopting Smarter Technology Today for Major Functions	206
Managing Law Enforcement Issues Cognitively	207
The Problem of Correlating Crime Data	207
The COPLink Project	208
Smart Energy Management: From Visualization to Distribution	209
The Problem of Integrating Regional Utilities Management	209
The Area Energy Management Solutions Project	209
The Cognitive Computing Opportunity	210
Protecting the Power Grid with Machine Learning	211
The Problem of Identifying Threats from New Patterns	211

The Grid Cybersecurity Analytics Project	211
The Cognitive Computing Opportunity	211
Improving Public Health with Cognitive Community Services	212
Smarter Approaches to Preventative Healthcare	212
The Town Health Station Project	212
The Cognitive Computing Opportunity	213
Building a Smarter Transportation Infrastructure	213
Managing Traffic in Growing Cities	213
The Adaptive Traffic Signals Controller Project	214
The Cognitive Computing Opportunity	214
Using Analytics to Close the Workforce Skills Gap	215
Identifying Emerging Skills Requirements and Just-in-Time Training	215
The Digital On-Ramps (DOR) Project	215
The Cognitive Computing Opportunity	216
Creating a Cognitive Community Infrastructure	217
The Smart + Connected Communities Initiative	217
The Cognitive Computing Opportunity	218
The Next Phase of Cognitive Cities	218
Summary	219
Chapter 13 Emerging Cognitive Computing Areas	221
Characteristics of Ideal Markets for Cognitive Computing	222
Vertical Markets and Industries	223
Retail	224
Cognitive Computing Opportunities	224
Retail Staff Training and Support	226
Travel	226
Cognitive Computing Opportunities for the Travel Industry	227
Transportation and Logistics	228
Cognitive Computing Opportunities for Transportation and Logistics	228
Telecommunications	229
Cognitive Computing Opportunities for Telecommunications	229
Security and Threat Detection	230
Cognitive Computing Opportunities for Security and Threat Detection	230
Other Areas That Are Impacted by a Cognitive Approach	231
Call Centers	231
Cognitive Computing Opportunities	231
Solutions in Other Areas	232
Summary	233

Chapter 14	Future Applications for Cognitive Computing	235
	Requirements for the Next Generation	236
	Leveraging Cognitive Computing to Improve Predictability	236
	The New Life Cycle for Knowledge Management	236
	Creating Intuitive Human-to-Machine Interfaces	237
	Requirements to Increase the Packaging of Best Practices	238
	Technical Advancements That Will Change the Future of Cognitive Computing	239
	What the Future Will Look Like	239
	The Next Five Years	240
	Looking at the Long Term	241
	Emerging Innovations	242
	Deep QA and Hypothesis Generation	243
	NLP	243
	Cognitive Training Tools	244
	Data Integration and Representation	245
	Emerging Hardware Architectures	245
	Neurosynaptic Architectures	246
	Quantum Architectures	248
	Alternative Models for Natural Cognitive Models	249
	Summary	249
	Glossary	251
	Index	261