

Index

A

accessibility algorithm 241, 248
 active data cube browser 266
 active rule and data warehouse integration 254
 active rule manager 265
 active rule schema metadata 266
 additivity 35
 advanced modeling 30
 aggregation 35
 availability window (AW) 233

B

business-activity monitoring (BAM) 88
 business process optimization (BPO) 92

C

common log format (CLF) 11
 conceptual design 36
 conceptual modeling 25
 convergence 31
 cross-dimension attribute 31
 cube 290

D

data-driven method (DDM) 10
 data availability 253
 database management systems (DBMS) 283
 database management systems (DBMSs) 302
 data cleaning v–vii, ix–xv, 43–58
 data comparability 253
 data cube manager 264
 data marts 100, 104, 172
 data warehouse (DW) 284
 data warehouse active rule syntax 259

data warehouse design ix, 24
 data warehouse engineering benchmark (DWEB) 302
 data warehouse quality metric 68
 data warehouses (DWs) 24
 DBMS (database management systems) 283
 decisional model 37
 decision support systems (DSS) 170, 171, 174, 194
 descriptive attribute 30
 designed approach 274
 DFM (see dimensional fact model) 26
 dimensional fact model (DFM) 26
 dimension attribute 29
 drill-down functions 325
 DW (data warehouse) 284
 DWs (see data warehouses) 24
 dynamic aspects 28
 dynamic dimension 324, 325, 326, 328, 329, 330, 331, 332, 333, 334, 335, 336
 dynamic hierarchies 34
 dynamicity 42, 329
 dynamic queries 72, 87

E

ELDA (evaluation datawarehouse quality) 61, 74
 entity/relationship (E/R) model 25
 entry points 30
 European social fund (ESF) 21
 extract, transform, and load (ETL) 302
 extraction, transformation, and loading (ETL) tool 90
 extraction time (ET) 233

F

factEntities 194, 196, 197, 198, 202, 204,
206, 208, 209, 210, 211, 221, 223, 2
24, 225
FactEntity model 228
factEntity model 196, 198, 199, 200
feature instances 285
flexible entity resolution architecture 46
flow measures 35
fuzzy duplicates 44, 45, 46, 47, 56
fuzzy if-then rules 48, 51

G

g-ordering 275
generation algorithm 309, 313
geographical information systems (GIS) 283
geometric spatial granularity 213, 214, 215
goal driven method (GDM) 15
Gorder process 275, 279
gorder query service 275
granularity (Gr) 234

I

integrating XML 158, 171
integration methodology 232, 234, 240
ipercube 69, 70, 71, 72, 73, 74, 76, 77, 78
, 83, 84
ISD goals 6

J

Java development kit (JDK) 264

K

key performance indicators (KPI) 92
k nearest neighbor (kNN) 275

L

LOG extraction 246

M

M-OLAP architecture nodes 152
machine learning, and data cleaning v–vii, ix–
xv, 43–58
MADS model 227

map cube operator 287
membership functions (MFs) 48
metadata repositories 265
metamodel 3, 6, 16, 17, 19, 22, 197, 228,
306, 307, 321, 322
metrics 30
mobile objects simulator 275
multi-level spatial measure (MuSD) 289
multidimensional databases 40, 42, 156, 157,
226, 228, 229, 299
multidimensional pattern 172, 173, 176, 187
, 189, 190
multiple arc 32

N

NodeCube 147

O

OLAP dimensions 155
online analytical processing (OLAP) 24
online transaction processing (OLTP) 24
optional arc 31
optional arcs 31
organizational model 37
OutBook 147

P

parameterization 307, 311, 316
particle swarm optimization 134, 135, 140, 1
41, 143, 144, 145
particle swarm optimization (PSO) 134
pattern pre-instantiation 182
periodic data 237
Pilote system SDW(l/t) 279
product-order-supplier model 304
product dimension 221, 330, 331, 335

Q

QualifQuerySCube 147

R

ragged (or incomplete) hierarchy 33
ragged hierarchies 33
raw textual measure 161, 164
read skew anomaly 324, 326, 327, 328, 336

Index

refresh algorithms 317
related data cubes 268
relational OLAP) 164, 306
RWE-graph 186, 187, 188

S

salary attribute 333, 334
scale factor (SF) 303
schema 290
secondary event 30
semantic granularity 201, 218
semi-periodic data 237
shared hierarchies 33
simple spatial features 286
simulation run periods 268
simulation times 268
SkewT diagram 245, 247
slowly changing dimensions
 324, 325, 326, 328, 336
snapshot data 237
sorted neighborhood method (SNM) 46
spatial data 283, 284
spatial data warehousing 282
spatial fact 289
spatial location and telemetric data warehouse
 (SDW(l/t) 274
spatial measure levels 292
spatial multidimensional data model 285
spatial multidimensional models 286
spatial OLAP (SOLAP) 284
spatial telemetric data warehouse (STDW) 274
star schema 25
star schema data warehouse 324
star schema metadata 266
StateDistrib 147
stock measures 35
storage time (ST) 234

T

temporal-based Data 254, 255, 256, 257, 258
temporal hierarchy 33
temporally-based data consistency mechanism
 261
temporal properties of data 238
temporal requirements algorithms 243
TPC-D 303, 304, 305, 319, 322
TPC-DS 303, 304, 305, 306, 311, 316, 318
 , 321, 322
TPC-H 303, 304, 305, 319
transaction processing performance council
 (TPC) 130, 154, 302, 303
transaction time (TT) 234
transient data 237
tree-dimensional workflow 93

U

unbalanced (or recursive) hierarchy 34
unbalanced hierarchies 34
unit measures 35
user driven method (UDM) 6
US National Weather Service Web site 245

V

valid interval temporal 255, 257, 268
virtual factEntities 194, 197, 198, 202, 204
 , 206, 208, 209, 210, 211, 223, 224,
 225

W

warehousing XML 158
WebCT 10, 11, 14
workflow management (WfM) 89

X

XML documents 158, 159, 160, 164, 169