

Jochen Müller

Transformation operativer Daten zur Nutzung im Data Warehouse

**Mit einem Geleitwort
von Prof. Dr. Roland Gabriel**



Deutscher Universitäts-Verlag

Inhaltsverzeichnis

Abbildungsverzeichnis	XIII
Abkürzungsverzeichnis	XV
1 Einleitung	1
1.1 Problemstellung und Zielsetzung	2
1.2 Vorgehensweise	3
1.3 Begriffsbestimmungen	4
1.3.1 Wissen, Information und Daten	5
1.3.2 Informationssystem	8
1.3.3 Kategorien betrieblicher Informationssysteme	10
1.3.4 Qualität von Daten	14
2 Operative Informationssysteme	19
2.1 Bedeutung und Ziele	20
2.1.1 Historie	20
2.1.2 Nutzenaspekte transaktionsorientierter Informationssysteme im Unternehmen	24
2.1.3 Vorgehensmodelle zur Entwicklung von datenbankbasierten Informationssystemen	26
2.2 Datenmodelle und Repräsentationsformen	32
2.2.1 Datenmodell	32
2.2.2 Semantische Datenmodelle	35
2.2.3 Logische Datenmodelle	37
2.2.3.1 Relationale Modelle	37
2.2.3.2 Alternative Repräsentationsformen	42
2.3 Architekturen und Komponenten	44
2.3.1 Anforderungen an Datenbanksysteme	47
2.3.2 Klassifikationsmerkmale für Datenbanksystemarchitekturen	50
2.3.3 Architekturarten	53
2.3.3.1 Zentrale Datenbanksysteme	53
2.3.3.2 Verteilte Datenbanksysteme	59
2.3.3.3 Multidatenbanksysteme	67

2.4	Metadatenverwaltung transaktionsorientierter Informationssysteme	70
2.4.1	Arten von Metadaten für transaktionsorientierte Informationssysteme	71
2.4.2	Verwaltungssysteme für Metadaten	72
2.4.3	Stellenwert von Metadaten im Rahmen der operativen Informationsverarbeitung	75
2.5	Datenaustausch auf operativer Ebene	78
3	Analyseorientierte Informationssysteme	81
3.1	Bedeutung und Ziele	82
3.1.1	Stellenwert analyseorientierter Informationssysteme im Unternehmen	83
3.1.2	Charakteristika des Data Warehouse-Konzepts	86
3.2	Datenmodelle und Repräsentationsformen	91
3.2.1	Mehrdimensionale Informationsstrukturen	92
3.2.2	Repräsentationsformen auf relationaler Basis	95
3.2.3	Repräsentationsformen auf multidimensionaler Basis	99
3.3	Architekturen und Komponenten	102
3.3.1	Data Warehouse	104
3.3.2	DataMart	114
3.3.3	Endbenutzerwerkzeuge	117
3.4	Metadatenverwaltung analyseorientierter Informationssysteme	120
3.4.1	Arten von Metadaten für analyseorientierte Informationssysteme	121
3.4.1.1	Strukturorientierte Metadaten	122
3.4.1.2	Funktionsorientierte Metadaten	128
3.4.1.3	Versionierung der Metadaten	130
3.4.2	Stellenwert von Metadaten im Umfeld analyseorientierter Informationssysteme	133
3.4.3	Verwaltungssysteme für Metadaten	136
3.5	Schnittstellen zu den Vorsystemen	138
4	Funktionale und inhaltliche Aspekte der Transformation operativer Daten für analyseorientierte Informationssysteme	143
4.1	Einordnung des Problembereichs	146
4.2	Anwendungsfälle einer Datentransformation für das Data Warehouse	149
4.2.1	Erzeugung der Data Warehouse-Datenbestände	150

Inhaltsverzeichnis

4.2.2	Aktualisierung der Data Warehouse-Datenbestände	151
4.3	Gewinnung und Transformation von Daten für das Data Warehouse	152
4.3.1	Aktivitäten auf Seiten des operativen Systems	154
4.3.1.1	Zeitstempel-gesteuerte Verfahren	156
4.3.1.2	Datenübergabe durch Anwendungsprogramme	158
4.3.1.3	Protokollierung relevanter Datenbank-Transaktionen	160
4.3.1.4	Auswertung der Logdateien	164
4.3.1.5	Schnappschuß-Vergleichsverfahren	165
4.3.1.6	Unterstützung durch Werkzeuge	166
4.3.2	Transformation im engeren Sinne	167
4.3.2.1	Basisfunktionen zur Durchführung der Transformationsschritte	169
4.3.2.2	Beseitigung syntaktischer Heterogenität	172
4.3.2.3	Datenaufbereitung	175
4.3.2.3.1	Datenreinigung	176
4.3.2.3.2	Entfernung irrelevanter Datenobjekte	178
4.3.2.3.3	Behandlung fehlender Werte	178
4.3.2.4	Modellkonvertierung	182
4.3.2.5	Homogenisierung des Datenbestands	186
4.3.2.6	Gesamtsicht	190
4.3.3	Aktivitäten auf Seiten der Zielumgebung	191
4.3.3.1	Datenübernahme	192
4.3.3.2	Erzeugen abgeleiteter Werte	196
4.4	Datentransformation zwischen Komponenten des analyseorientierten Informationssystems	198
4.5	Schlußfolgerungen	204
5	Anforderungen an die Transformationskomponente im Rahmen der Entwicklung und Nutzung eines analyseorientierten Informationssystems	207
5.1	Besonderheiten bei der Entwicklung von Komponenten für analyseorientierte Informationssysteme	208
5.2	Vorgehen bei der Data Warehouse-Entwicklung	211
5.2.1	Planung und Entwurf	212
5.2.1.1	Umfeldanalyse	213
5.2.1.2	Architekturplanung	215

Abbildung 23:	Drei-Schichten-Modell der Transformationskomponente	153
Abbildung 24:	Extraktion von Detail-Datensätzen über Zeitstempel am Master-Datensatz	157
Abbildung 25:	Datenübergabe durch das Anwendungsprogramm	159
Abbildung 26:	Protokollierung von Transaktionen mit Hilfe von Datenbanktriggern	163
Abbildung 27:	Beispiele syntaktischer Umwandlung von Datenbeständen	173
Abbildung 28:	Schritte zur Datenaufbereitung	176
Abbildung 29:	Modellkonvertierung	183
Abbildung 30:	Modellkonsolidierung von Datenbeständen unterschiedlicher Quellen	185
Abbildung 31:	Inhaltliche Homogenisierung von Datenbeständen	188
Abbildung 32:	Gesamtsicht des Umwandlungsprozesses	191
Abbildung 33:	Alternative Konzepte zur Data Mart-Population	202
Abbildung 34:	Vorgehen im Rahmen von Data Warehouse-Projekten	224
Abbildung 35:	Data Warehouse-Lebenszyklus	228