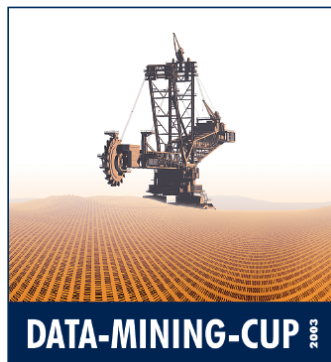


---

**Vortrag auf den 4. DATA-MINING-CUP Anwendertagen  
(Chemnitz, 18.-20. Juni 2003)**



<http://www.data-mining-cup.de>

Copyright-Hinweis:

Das Urheberrecht des folgenden Vortrags liegt beim Author. Verbreitung, Vervielfältigung und Kopie, auch auszugsweise, ist nur mit schriftlicher Genehmigung des Authors erlaubt.

# Das Data Warehouse - zentrale Drehscheibe für Information

aufgezeigt an Beispielen aus dem CRM bei Quelle

QUELLE AG  
Database Management, Data Mining  
Dr. Gernot Schreiber  
Jakobinenstr. 5-7  
90762 Fürth  
Tel. +49 (0)911 14-22084  
Mail: gernot.schreib@quelle.de

## Data Warehouse – Zentrale Drehscheibe für Information

### Gliederung

#### Granulares Data Warehouse

- Basis für Wissensgenerierung und CRM-Kampagnen
- Inhalt und Struktur
- Zentrum wesentlicher Prozesse
- Quelle-Data Warehouse

#### Dynamische Data Marts

- ETL-Tool
- Aggregationsstufen/Vorverdichtungen
- Merkmalsdatenbank

#### CRM-Controlling mit dynamischen Data Marts

- Wissensgenerierung
- Einsatz und Erfolgskontrolle

#### Journalisierung

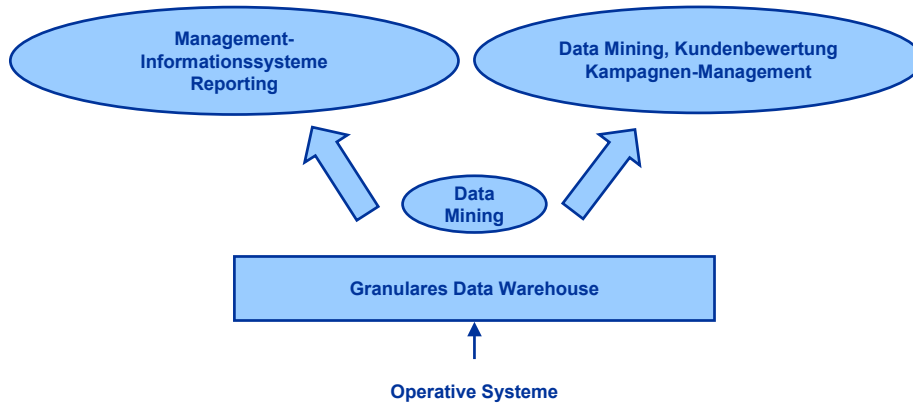
#### Anwendung

#### Zusammenfassung

## Data Warehouse – Zentrale Drehscheibe für Information

## Data Warehouse und Data Mining bilden die Basis für Wissensgenerierung und Umsetzung von CRM-Kampagnen

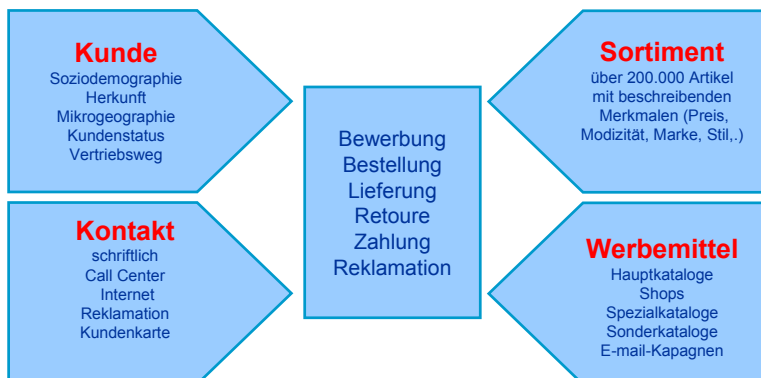
QUELLE



## Data Warehouse – Zentrale Drehscheibe für Information

## Granulares Data Warehouse - Inhalt und Struktur

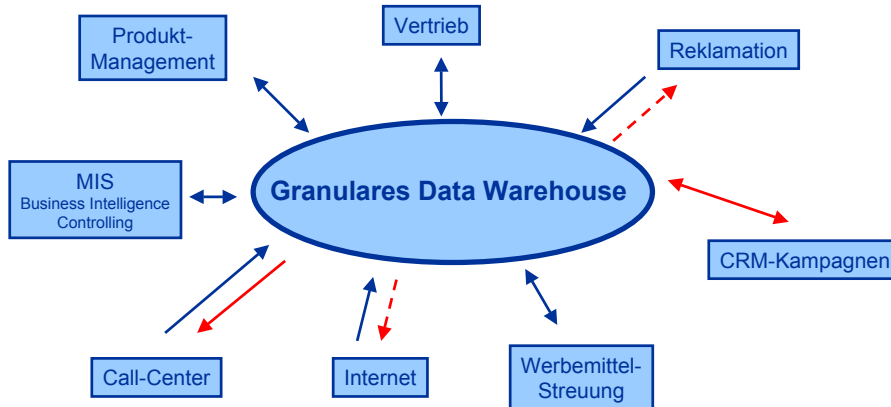
QUELLE



Kunden- und Artikel-Stammdaten  
 Verhaltensdaten auf granularer Ebene  
 Historische Sichten (Journalisierung der statischen Daten)

## Data Warehouse – Zentrale Drehscheibe für Information

## Das Data Warehouse integriert die wesentlichen Prozesse



➔ Sammlung, Konsolidierung und Speicherung der Daten aus allen kundenbezogenen Transaktionen

← Bereitstellung von zielgerichteten Informationen für die Steuerung der CRM-Prozesse

## Data Warehouse – Zentrale Drehscheibe für Information

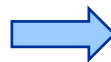
## Granulares Data Warehouse erfordert ETL-Tool

Granulare Daten, d. h. Informationen auf der feinsten Ebene wie z. B. Bestell- oder Kaufpositionen bieten:

- **einerseits** eine *große Flexibilität* bei der Berechnung von Merkmalen können
- **andererseits** nicht *unmittelbar* in Typologien oder Selektionen übernommen werden.

**Ziel:** Unter Erhaltung der Flexibilität

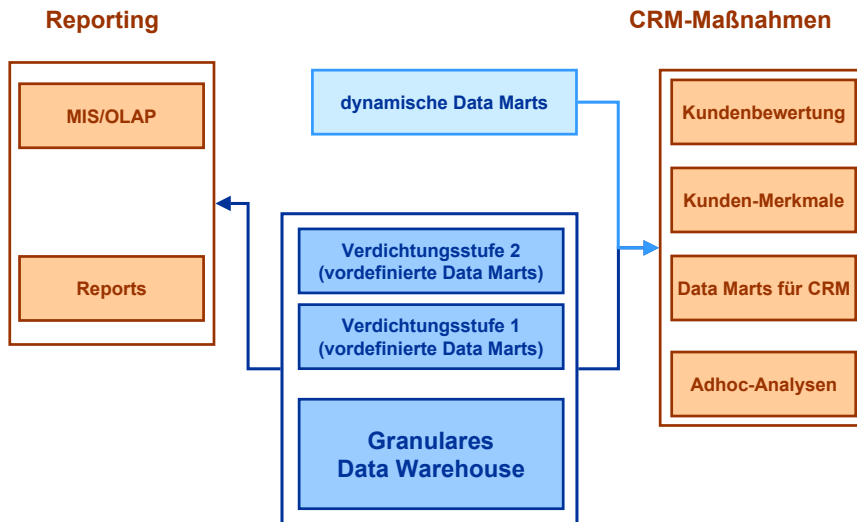
- Standardisierung
- Sichern der Korrektheit
- Vereinfachung
- Dokumentation und Verwaltung



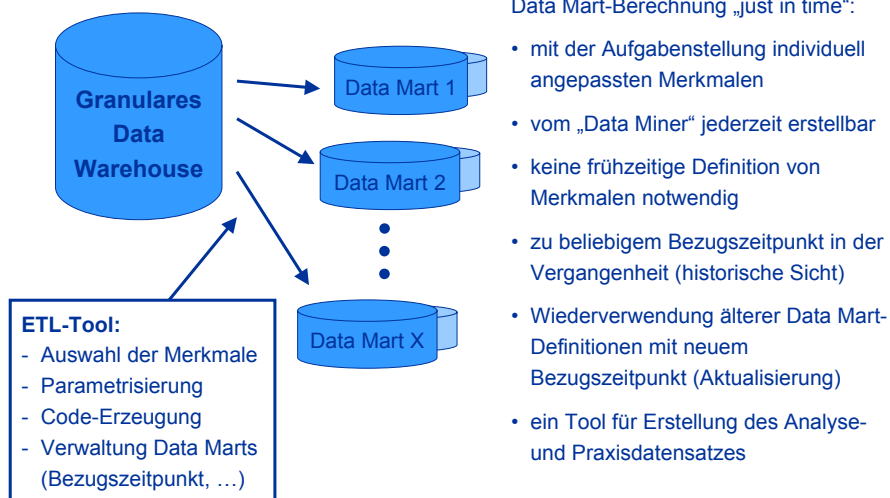
**ETL-Tool**

der Berechnung der notwendigen Merkmale und Verdichtungen.

## Aggregationsstufen (Vorverdichtungen)



## Dynamische Data Marts



Data Mart-Berechnung „just in time“:

- mit der Aufgabenstellung individuell angepassten Merkmalen
- vom „Data Miner“ jederzeit erstellbar
- keine frühzeitige Definition von Merkmalen notwendig
- zu beliebigem Bezugszeitpunkt in der Vergangenheit (historische Sicht)
- Wiederverwendung älterer Data Mart-Definitionen mit neuem Bezugszeitpunkt (Aktualisierung)
- ein Tool für Erstellung des Analyse- und Praxisdatensatzes

## ETL-Tool: Merkmals-Auswahl

HA-TAIFUN @dbih - [Laufvariablen definieren]

Systemparameter Abgleichstabellen Datenerstellung Edit Block Field Record Query Window Help

DBM>Data Mining

DATENERSTELLUNGSLAUF

Laufbezeichnung

Demo - Scorekarte

Variablengruppen auflösen Variablen kopieren (von anderem Erstellungslauf)

MÖGLICHE VARIABLEN

Variablenbezeichnung	Typ
HKH BH /Haushalt NH von-bis VG	VARG
HKH BH /Haushalt NH,PK von-bis	EVAR
HKH BH /Haushalt NH,PK von-bis VG	VARG
HKH BH-Anteil /Hsh. BH v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. BH v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. PK v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. PK v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. S0 v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. S0 v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. S0 v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. S0 v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. SB v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. SB v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. SC v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. SC v. Gesamt von-bis VG	VARG
HKH BH-Anteil /Hsh. HH v. Gesamt von-bis	EVAR
HKH BH-Anteil /Hsh. HH v. Gesamt von-bis VG	VARG

LAUFVARIABLEN

Variablenbezeichnung
HEH BH /Haushalt S0 von-bis
HEH Ersta. bestellt /Hsh.
HEH Letztn.bestellt /Hsh. NH
HPK Anzahl Personen in Haushalt
HPK Alter: aelteste Person in Haushalt

AKTUALPARAMETER

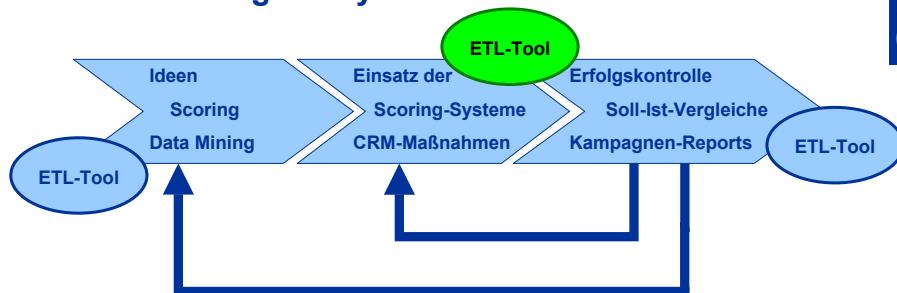
Beschreibung	Wert
Eigner/-kombination	Quelle AG
A-Sortiment/-Cluster	SKA 01 Damen-Oberbekleidung
Datensuntergrenze: von	-24 Monate
Datensobergrenze: bis	Bezugsdatum

Bezeichnung der Variable

Record: 1/5

Record: 1/5

## CRM-Controlling mit dynamischen Data Marts

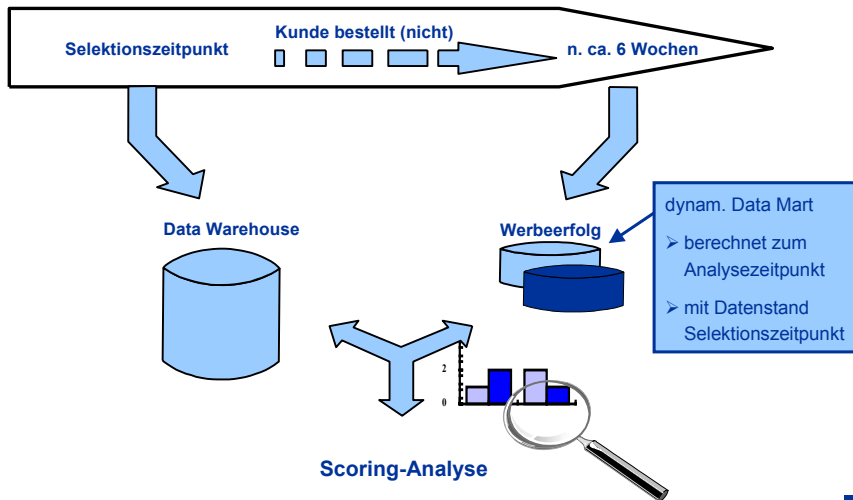


## Closed-Loop:

- Eingesetzte Scoring-Systeme werden regelmäßig überprüft
- Neue Marketing-Konzepte bzw. alternative Strategien werden getestet
- Das Gelernte hat Einfluss auf zukünftige Entscheidungen

## Response-Prognose (Analyse-Ansatz)

Einsatz einer Lernstichprobe für einen Katalog



## Zusammenfassung

### Granulares, journalisierendes Data Warehouse

- ⇒ realisiert eine Abbildung der Realität
- ⇒ bevorzugt keine aktionsgetriebene Sichtweise
- ⇒ erlaubt Berechnung von Data Marts auch für Bezugszeitpunkte in der Vergangenheit
- ⇒ steht im Zentrum aller wesentlichen Prozesse

### Dynamische Data Marts

- ⇒ ergänzen das DWH
- ⇒ erlauben aktionsspezifische Aggregationsstufen und Verdichtungen
- ⇒ werden vom „Data Miner“ zeitnah angelegt
- ⇒ sichern (auf einer Merkmalsdatenbank basierend) Korrektheit, Wiederverwendbarkeit und Dokumentation