

Fragestellungen

- Was ist die Aufgabe von InfoAreas im SAP BW System?
- Für welche InfoObjekt Typen können jeweils InfoObjekt Kataloge angelegt werden?
- Können identische InfoObjekte in mehreren InfoObjekt Katalogen aufgenommen werden?
- Können InfoObjekt Kataloge denselben technischen Namen verwenden?
- Was denken Sie? In welcher InfoArea befinden sich InfoObjekte, wenn diese noch keinem InfoObjekt Katalog zugeordnet wurden? Funktioniert dies überhaupt?
- Wie fügt man bereits angelegte InfoObjekte nachträglich in InfoObjekt Kataloge hinzu?
- Erklären Sie, wie InfoObjekte definiert sind.
- Welche Typen von InfoObjekten können Sie unterscheiden?
- Was ist der Unterschied zwischen einem Navigations-Attribut und einem Anzeige-Attribut?
- Könnte eine Kennzahl selbst ein Attribut sein oder gar ein Navigationsattribut?
- Wie lautet der Schlüssel bei Verwendung eines Navigations-Attributs in Bezug zum führenden InfoObjekt? Schreiben Sie ein Beispiel mit dem Navigationsattribut SS6XXI12 für das ProfitCenter auf.
- Was bedeutet die Klammerung von InfoObjekten?
- Können Hierarchien versionsabhängig sein?
- Wo kann in der InfoObjekt-Pflege für Merkmale angesteuert werden, ob InfoObjekte berechtigungsrelevant sind?
- Im Bereich Allgemein in der InfoObjekt Pflege kann eingestellt werden, dass das InfoObjekt ausschließlich Attribut ist. Was bedeutet diese Einstellung?
- Für welche Datentypen können DataSources angelegt werden?

- Welche Feldlängen können die Bezeichnungen der DataSource für Kurz-, Mittel- und Lang-Texte maximal haben?
- Warum ist das InfoObjekt Kostenrechnungskreis zwingend mit als Feld in die DataSource aufzunehmen?
- Ist die Positionierung der Felder in der DataSource für das Hochladen von Daten einzuhalten? Wenn ja, warum?
- Bevor Sie den Datenfluss zu den InfoObjekten aufbauen können, müssen die Merkmale als InfoProvider eingefügt werden. Nehmen Sie dies bitte für die Objekte Kostenstelle und ProfitCenter entsprechend vor (Tipp: Schauen Sie in der Data Warehousing Workbench in der Sicht der InfoProvider nach). Für welche Art von Stammdaten könnten danach Datenflüsse für die InfoObjekte erzeugt werden und an welcher Stelle haben Sie bei der Definition der InfoObjekte dafür gesorgt?
- Wie ist ein Standard-DSO technisch aufgebaut? Beschreiben Sie es.
- Was passiert bei der Aktivierung eines Requests im DSO?
- Welche DataStore-Objektypen unterscheidet man?
- Wie unterscheidet sich ein schreiboptimiertes DSO von einem Standard-DSO in dessen technischem Aufbau?
- Ist das Reporting auf einem DSO performanter als das Reporting auf einem Standard-Basis InfoCube?
- Wie ist das erweiterte Sternschema aufgebaut? Skizzieren Sie es aus technischer Sicht. Beantworten Sie noch die Frage, welche Vorteile das erweiterte Sternschema bezüglich der Verwendung von Stammdaten beinhaltet.
- Aus wie vielen Dimensionen kann ein InfoCube maximal bestehen?
- Wie viele Merkmalswerte können in einer Dimension eingesteuert werden?
- Wie viele Kennzahlen kann ein InfoCube insgesamt aufnehmen?
- Was bedeutet es, eine Dimension als Line-Item-Dimension zu führen?

- Was genau ist ein MultiProvider? Beschreiben Sie dieses Objekt.
- Aus welchen Kombinationen von InfoProvider kann sich ein MultiProvider zusammensetzen?

Was ist Business Intelligence?

Die Aufgabe von Business Intelligence

Unternehmen erzeugen bei sämtlichen Geschäftsaktivitäten Daten. Auf Basis dieser Daten werden in allen Unternehmensbereichen von Mitarbeitern aller Ebenen Entscheidungen getroffen. Business Intelligence (BI) hat die Aufgabe, die große Menge an Unternehmensdaten gezielt zu bündeln und aufzubereiten. Durch Auswertung der Daten mit BI-Werkzeugen lassen sich Erkenntnisse gewinnen, die den Entscheidungsprozess im Unternehmen unterstützen. Eine schnelle Berichterstattung über Ablauf und Ergebnisse der Geschäftsprozesse, Analyse und Interpretation von Daten über Kunden, Lieferanten und interne Vorgänge bis hin zu dynamischer Planung wird ermöglicht. Somit hilft Business Intelligence Geschäftsprozesse zu optimieren, marktgerecht und schnell zu agieren, und schafft somit entscheidende Wettbewerbsvorteile Ihres Unternehmens.

Schlüsselbereiche von Business Intelligence

Eine vollständige Business-Intelligence-Lösung untergliedert sich in verschiedene Bereiche. **SAP NetWeaver Business Intelligence**(SAP NetWeaver BI) bietet für alle diese Bereiche umfassende Werkzeuge, Funktionalitäten und Prozesse:

Ein Data Warehouse sorgt für **Integration, Ablage und Verwaltung** von Unternehmensdaten aus den verschiedensten Quellen.

Wenn im Data Warehouse eine integrierte Sicht auf die relevanten Daten vorliegt, können **Analyse und Planung** erfolgen. Um aus den Daten die entscheidenden Erkenntnisse zur Verbesserung Ihrer Geschäftsabläufe zu gewinnen, liefert SAP NetWeaver BI Methoden für die multidimensionale Analyse, d.h. betriebswirtschaftliche Kennzahlen wie z.B. Absatzmengen oder Umsatz können anhand unterschiedlicher Bezugsobjekte wie z.B. Produkt, Kunde, Zeit analysiert werden. Des Weiteren stehen Methoden zur Mustererkennung im Datenbestand (Data Mining) zur Verfügung. Zudem ermöglicht SAP NetWeaver BI die Durchführung von Planung auf der Basis der Daten im Data Warehouse.

Werkzeuge für den Zugriff und die Visualisierung ermöglichen die Darstellung der gewonnenen Erkenntnisse, die Analyse und die Planung der Unternehmensdaten in verschiedenen Detailgraden und Arbeitsumgebungen (Web, Microsoft Excel).

Durch **Publizieren von BI-Inhalten** können Sie die Informationen an die verschiedenen in die Entscheidungsprozesse eingebundenen Mitarbeiter in Ihrem Unternehmen flexibel verteilen, z.B. per E-Mail oder über ein Unternehmensportal.

Neben diesen Bereichen spielen **Performance** und **Sicherheit** eine wichtige Rolle, wenn es darum geht, die entscheidungsrelevanten Informationen zur richtigen Zeit den richtigen Mitarbeitern zur Verfügung zu stellen.

Vorkonfigurierte Informationsmodelle in Form des so genannten **BI Content** ermöglichen eine effiziente und kostengünstige Einführung von SAP NetWeaver BI.

Integration, Ablage und Verwaltung der Daten

Umfassende, aussagekräftige Datenanalysen werden erst möglich, wenn der meist in unterschiedlichen Formaten und Quellen vorliegende Datenbestand zu einer betriebswirtschaftlichen Fragestellung gebündelt und integriert wird. Die Grundlage einer Business-Intelligence-Lösung bildet daher das Data Warehouse.

Im Enterprise Data Warehouse von SAP NetWeaver BI werden Unternehmensdaten zentral gesammelt. Die Daten werden i. d. R. aus verschiedenen Quellen ausgelesen (extrahiert) und ins SAP NetWeaver BI geladen. Dabei unterstützt SAP NetWeaver BI Quellen beliebiger Art -relationale wie multidimensionale, SAP- wie Nicht-SAP-Quellen. Anschließend werden technische Bereinigungen durchgeführt und betriebswirtschaftliche Regeln angewendet, um die Daten für Auswertungen zu konsolidieren. Die konsolidierten Daten werden dann im Enterprise Data Warehouse gespeichert. Dieser Gesamtprozess wird Extraktion, Transformation und Laden (ETL) genannt.

Die Datenablage im Enterprise Data Warehouse kann je nach Anforderung in unterschiedlicher Granularität in verschiedenen Schichten der Data-Warehouse-Architektur erfolgen. Der Datenfluss beschreibt den Weg, den die Daten durch die Data-Warehouse-Schichten zurücklegen, bis sie bereit für die Auswertung sind.

Die Verwaltung der Daten im Enterprise Data Warehouse umfasst u.a. die Steuerung der Prozesse, welche die Daten in das Enterprise Data Warehouse übertragen und innerhalb des Enterprise Data Warehouse verteilen, wie auch die Umsetzung von Strategien zur optimalen Datenhaltung und Historienführung (Begrenzung des Datenvolumens), auch Information Lifecycle Management genannt.

Durch Extraktion in nachgelagerte Systeme können Sie die im Enterprise Data Warehouse konsolidierten Daten weiteren BI-Systemen oder weiteren Anwendungen in Ihrer Systemlandschaft zur Verfügung stellen.

Ein Metadaten-Konzept ermöglicht, die Daten in SAP NetWeaver BI mit Hilfe von Definitionen oder Informationen in strukturierter und unstrukturierter Form zu dokumentieren.

Die Data Warehousing Workbench ist die zentrale Arbeitsumgebung, welche die Werkzeuge zur Durchführung der Aufgaben im Enterprise Data Warehouse von SAP NetWeaver BI zur Verfügung stellt.

Extraktion, Transformation und Laden (ETL)

SAP NetWeaver BI bietet flexible Möglichkeiten, Daten aus unterschiedlichen Quellen zu integrieren. In Abhängigkeit von Ihrer Data-Warehousing-Strategie für das jeweilige Anwendungsszenario können Sie die Daten aus der Quelle extrahieren und ins SAP-NetWeaver-BI-System laden oder direkt auf Daten in der Quelle zugreifen, ohne sie im Enterprise Data Warehouse physisch abzulegen. In diesem Fall erfolgt die Integration der Daten im Enterprise Data Warehouse virtuell.

Quellen für das Enterprise Data Warehouse können operationale, relationale Datenbestände (z.B. in SAP-Systemen), Dateien oder Altsysteme sein. Auch multidimensionale Quellen, d.h. Daten aus anderen BI-Systemen, sind möglich. Transformationen ermöglichen die technische Bereinigung und Konsolidierung der Daten in betriebswirtschaftlicher Hinsicht.

Extraktion und Laden

Extraktions- und Übertragungsprozesse in die Eingangsschicht des SAP NetWeaver BI sowie der direkte Zugriff auf Daten sind über verschiedene Schnittstellen möglich, je nach Herkunft und Format der Daten. Auf diese Weise ermöglicht SAP NetWeaver BI die Integration von relationalen sowie multidimensionalen Daten ebenso wie von SAP- und Nicht-SAP-Daten:

- **BI Service API** (BI Service Application Programming Interface)

Das BI Service API ermöglicht Extraktion und Direktzugriff auf Daten aus SAP-Systemen in standardisierter Form. Dies können SAP-Anwendungssysteme oder auch SAP-NetWeaver-BI-Systeme sein. Die Datenanforderung wird dabei aus dem SAP-NetWeaver-BI-System heraus gesteuert.

- **Dateischnittstelle**

Die Dateischnittstelle ermöglicht Extraktion und Direktzugriff auf Dateien, wie beispielsweise csv-Dateien. Die Datenanforderung wird dabei aus dem SAP-NetWeaver-BI-System heraus gesteuert.

Transformation

Unter Anwendung von **Transformationen** werden die über die genannten Schnittstellen geladenen Daten innerhalb des SAP-NetWeaver-BI-Systems in den Data-Warehouse-Schichten aus einem Quellformat in ein Zielformat überführt. Die Transformation ermöglicht, die Daten zu konsolidieren, zu bereinigen und zu integrieren und somit technisch und semantisch zu vereinheitlichen, um eine Auswertbarkeit zu ermöglichen. Dazu werden Regeln angeboten, die beliebige Komplexität bei der Transformation der Daten erlauben.

Die Funktionalität reicht dabei von einer 1:1-Zuordnung der Daten über die Anwendung von umfangreichen Funktionen in Formeln bis hin zur kundenspezifischen Programmierung von Transformationsregeln. So können Sie für die Transformation beispielsweise Formeln definieren, die die Funktionen der so genannten Transformationsbibliothek verwenden.

Zur Erstellung von Formeln werden Grundfunktionen (z.B. und, wenn, kleiner, größer), verschiedenste Funktionen für Zeichenketten (z.B. Werte in Großbuchstaben darstellen), Datumsfunktionen (z.B. Quartal aus Datum berechnen), mathematische Funktionen (z.B. Division, Exponentialfunktion) und andere angeboten.

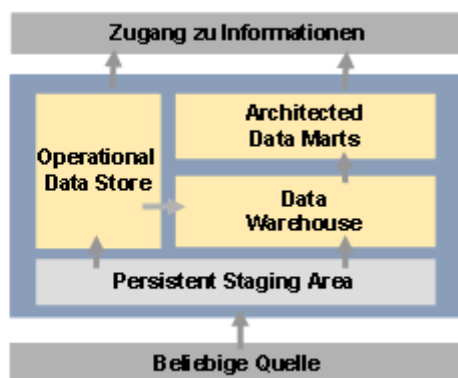
Datenablage und Datenfluss

SAP NetWeaver BI bietet umfangreiche Möglichkeiten zur Datenablage an. Dazu gehört die Implementierung eines Data Warehouses oder eines Operational Data Store wie auch die Schaffung der zur Analyse verwendeten Datenablagen.

Architektur

Eine mehrschichtige Architektur dient der Integration von Daten aus unterschiedlichen Quellen, der Transformation, Konsolidierung, Bereinigung und Ablage von Daten sowie der effizienten Bereitstellung der Daten zur Analyse und Interpretation. In den Schichten können die Daten in unterschiedlicher Granularität abgelegt werden.

Die folgende Grafik zeigt, welche Schichten das Data-Warehousing-Konzept von SAP NetWeaver BI umfasst:



• Persistent Staging Area (PSA)

Nach der Extraktion aus einer beliebigen Quelle gelangen die Daten in die Eingangsschicht des Enterprise Data Warehouses, die Persistent Staging Area (PSA). Die Daten aus dem Quellsystem werden in dieser Schicht unverändert abgelegt. Sie liefert den Backup-Status auf granularer Ebene, um später potentiell weitere Informationen zur Verfügung zu stellen, und sie ermöglicht im Fehlerfall ein schnelles Wiederaufsetzen.

• Data Warehouse

Der Weg der Daten vom PSA in die nächste Schicht dient qualitätssichernden Maßnahmen und den für eine einheitliche, integrierte Datensicht notwendigen Bereinigungen. In der Data-Warehouse-Schicht wird das Ergebnis dieser ersten Transformationen und Bereinigungen abgelegt. Sie bietet integrierte, granulare, historische, stabile Daten, die noch nicht für einen konkreten

Anwendungszweck verändert wurden, d.h. sie sind applikationsneutral. Das Data Warehouse bildet die Grundlage und zentrale Datenbasis für weitere (verdichtete) Datenhaltungen zu Analysezwecken (Data Marts). Ohne ein zentrales Data Warehouse können Erweiterung und Betrieb von Data Marts oft nur unzureichend gestaltet werden.

- **Architected Data Marts**

Die Data-Warehouse-Schicht beliefert die meist multidimensionalen Auswertungsstrukturen, die auch als Architected Data Marts bezeichnet werden. Dabei ist Data Mart nicht notwendig mit summiert bzw. aggregiert gleichzusetzen; auch hier findet man hoch granulare Strukturen, die sich aber allein an den Anforderungen der Auswertung orientieren.

- **DataStore Object (DSO)**

Ein DSO unterstützt die operative Datenanalyse. Die Daten werden in einem DSO fortwährend oder in kurzen Zeitabständen verarbeitet und für die operative Analyse gelesen. Somit kennzeichnet einen DSO eine hohe Datenaktualität auf meist unverdichteten Datenbeständen, wodurch operative Analysen optimal unterstützt werden.

Datenablage

Bei der Modellierung der Schichten stehen Ihnen zur physischen Ablage verschiedene Strukturen bzw. Objekte zur Verfügung, die je nach Anforderung verwendet werden können.

In der Persistent Staging Area (PSA) ist die Struktur der Quelldaten durch **DataSources** abgebildet. Zu einer DataSource werden die Daten einer betriebswirtschaftlichen Einheit (z.B. Kundenstammdaten oder Positionsdaten eines Auftrags) in einer transparenten, flachen Datenbanktabelle, der PSA-Tabelle, abgelegt. Die Datenablage in der Persistent Staging Area ist kurz- bis mittelfristig, sie liefert den Backup-Status für die nachfolgenden Datenablagen; daher sind Abfragen auf dieser Ebene und die Archivierung dieser Daten nicht möglich.

Während eine DataSource aus einer Menge von Feldern besteht, sind die im Datenfluss folgenden Datenablagen durch **InfoObjects** definiert. Die Felder der DataSource müssen über Transformationen in SAP NetWeaver BI den InfoObjects zugeordnet werden. InfoObjects sind somit die kleinsten (Metadaten-)Einheiten des BI. Durch sie werden die Informationen in strukturierter Form abgebildet, die zum Aufbau der Datenablagen benötigt werden. Sie untergliedern sich in **Kennzahlen, Merkmale** und **Einheiten**.

- Kennzahlen liefern die Bewegungsdaten, d.h. die Werte, die analysiert werden sollen. Dabei handelt es sich um Mengen, Beträge oder Stückzahlen, z.B. Absatzmengen oder Umsatzzahlen.

- Merkmale sind Ordnungsbegriffe wie z.B. Produkt, Kundengruppe, Geschäftsjahr, Periode oder Region. Sie geben Klassifizierungsmöglichkeiten des Datenbestands vor und sind damit Bezugsobjekte für die Kennzahlen. Merkmale können Stammdaten in Form von Attributen, Texten oder Hierarchien enthalten. Stammdaten sind Daten, die über einen längeren Zeitraum unverändert bleiben. Die Stammdaten einer Kostenstelle beispielsweise enthalten die Bezeichnung (Text), den Verantwortlichen (Attribut), den zugehörigen Hierarchiebereich (Hierarchie) usw.

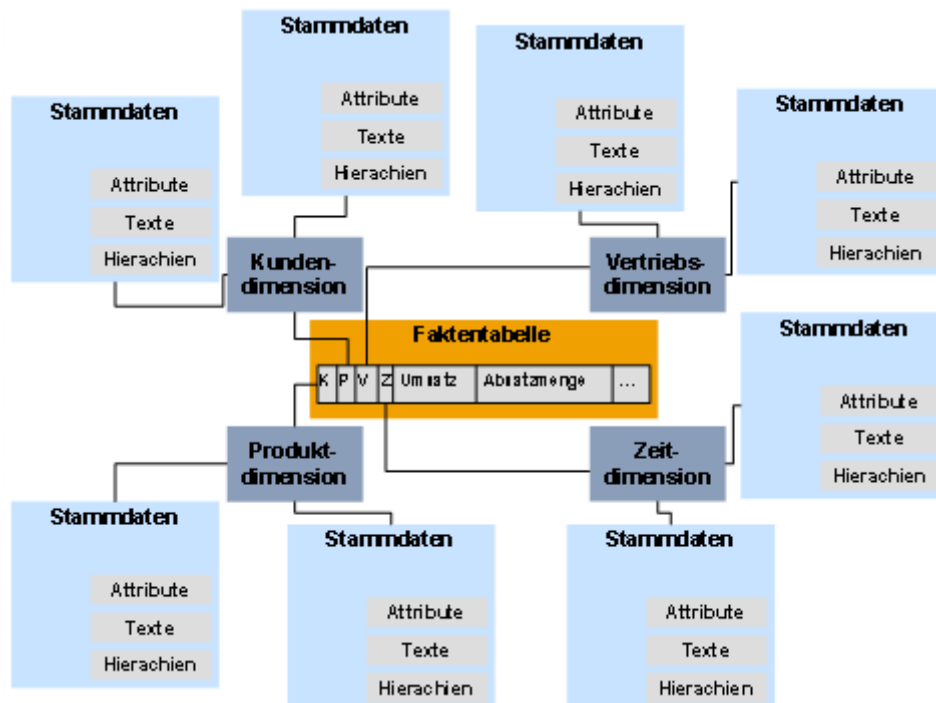
- Einheiten wie Währungen oder Mengeneinheiten geben den Werten der Kennzahlen den Kontext. Durch die durchgängige Nutzung von identischen InfoObjects zur Definition der Datenablagen in den unterschiedlichen Schichten wird Konsistenz auf Ebene der Metadaten gesichert.

DataStore-Objekte erlauben es, die Daten granular (Belegebene) und historisch vollständig zu speichern. Die Datenablage erfolgt dabei wie bei DataSources in flachen Datenbanktabellen. Ein DataStore-Objekt besteht aus Schlüssel (z.B. Belegnummer, Position) und Datenbereich. Im Datenbereich können neben Kennzahlen (z.B. Bestellmenge) auch Merkmale (z.B. Auftragsstatus) enthalten sein. Neben der Aggregation der Daten ist es auch möglich, die Dateninhalte zu überschreiben, z.B. um die Statusänderungen des Auftrags abzubilden. Dies ist insbesondere im Kontext von belegnahen Strukturen wichtig.

Die Modellierung einer mehrdimensionalen Ablage wird durch **InfoCubes** realisiert. Ein InfoCube ist eine Menge von relationalen Tabellen, die nach einem erweiterten Sternschema zusammengestellt sind. Es gibt eine (große, viele Zeilen enthaltende) Faktentabelle, welche die Kennzahlen des InfoCube enthält, sowie mehrere sie umgebende (kleinere) Dimensionstabellen, in denen die Merkmale des InfoCube abgelegt sind. Die Merkmale stellen hierbei den Schlüssel für die Kennzahlen dar. Die Ablage der Daten in einem InfoCube ist additiv.

Bei Abfragen auf einen InfoCube werden, wenn nötig, die Fakten bzw. Kennzahlen automatisch aggregiert (Summation, Minimum oder Maximum). Die Dimensionen fassen sachlogisch zusammengehörige Merkmale zusammen - z.B. eine Kundendimension, die aus der Kundennummer, der Kundengruppe und den Stufen der Kundenhierarchie bestehen kann oder eine Produktdimension, die aus der Produktnummer, der Produktgruppe und der Marke bestehen kann. Die Merkmale wiederum zeigen auf die Stammdaten (Texte oder Attribute des Merkmals). Die Fakten sind die auszuwertenden Kennzahlen wie Umsätze oder Absatzmengen. Faktentabelle und Dimensionen sind über identifizierende abstrakte Nummern (Dimensions-IDs) miteinander verknüpft.

Hierdurch werden die Kennzahlen des InfoCubes auf die Merkmale der Dimension bezogen. Diese Art der Modellierung ist optimiert für eine effiziente Datenanalyse. Die folgende Grafik skizziert den Aufbau eines InfoCubes:



Auf den physischen Datenablagen in Form von InfoObjects, InfoCubes und DataStore-Objekten können Sie logische Sichten (**MultiProvider**, **InfoSets**) bilden, z.B. um Daten verschiedener Datenablagen für einen gemeinsamen Auswertungszweck zur Verfügung zu stellen. Die Verknüpfung wird dabei über die gemeinsamen InfoObjects der Datenablagen hergestellt.

Der Oberbegriff für die physischen Datenablagen und die logischen Sichten auf diese lautet **InfoProvider**. Die Aufgabe eines InfoProviders ist es, den Werkzeugen für Analyse, Reporting und Planung die Daten zweckoptimiert zur Verfügung zu stellen.

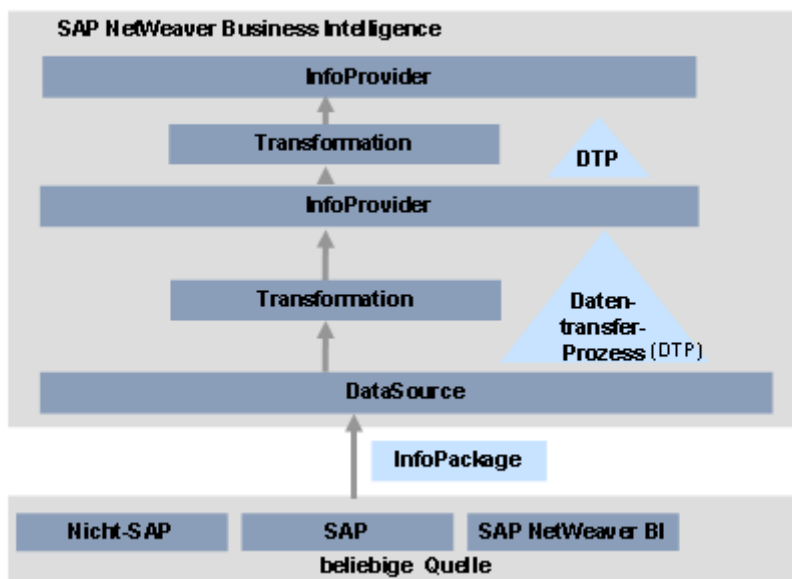
Datenfluss

Der Datenfluss im Enterprise Data Warehouse beschreibt, wie die Daten durch die Schichten geführt werden, bis sie letztendlich in der Form vorliegen, die dem Anwendungszweck genügt. Dadurch können Datenextraktion und -verteilung gezielt gesteuert werden sowie die Datenherkunft komplett aufgezeigt werden. Das Transferieren der Daten von einer Datenablage in die Nächste erfolgt mit Hilfe von Ladeprozessen. Für das Laden der Quelldaten in die Eingangsschicht des SAP NetWeaver BI, in die Persistent Staging Area, verwenden Sie das **InfoPackage**. Für das Laden der Daten innerhalb des BI von einer physischen Datenablage in die darauf folgende verwenden Sie den

Datentransferprozess (DTP) unter Anwendung der bereits beschriebenen Transformationsregeln. Dabei werden auch Zuordnungen zwischen den Feldern/InfoObjects der Quellablage und InfoObjects der Zielablage durchgeführt.

Einen Ladeprozess definieren Sie jeweils für eine Quell-Ziel-Kombination und legen hier das Bereitstellungsverfahren fest. Sie können für den Ladeprozess verschiedene Einstellungen vornehmen, teilweise abhängig von der Art der Daten und Quelle sowie Ziel der Daten. So können Sie beispielsweise Selektionen der Daten festlegen, einerseits um nur die relevanten Daten zu übertragen und andererseits um den Ladeprozess in Hinsicht auf Performance zu optimieren. Oder Sie können festlegen, ob der gesamte Datenbestand der Quelle geladen wird oder nur der seit dem letzten Ladezeitpunkt neu in der Quelle hinzugekommene Datenbestand. Letzteres bedeutet, dass Datentransferprozesse automatisch eine Deltaverarbeitung erlauben und dies für jedes Datenziel individuell. Bei InfoPackages, also dem Laden in das SAP-NetWeaver-BI-System, hängt die Verarbeitungsform (Delta oder gesamter Datenbestand) vom angesprochenen Extraktionsprogramm ab.

Die folgende Grafik zeigt, wie ein einfacher Datenfluss über zwei InfoProvider aussieht:



Data Warehousing Workbench:

Modellierung

Einsatzmöglichkeiten

Im Funktionsbereich Modellierung der Data Warehousing Workbench können Sie die BI-Objekte mit dem zugehörigen Datenfluss in Objektbäumen strukturiert anzeigen, Objekte neu anlegen, Anwendungen und Funktionen zu den Objekten aufrufen und den Datenfluss für die Objekte definieren.

Datenfluss im Data Warehouse

Der Datenfluss im Data Warehouse beschreibt, welche Objekte zur Designzeit und Prozesse zur Laufzeit benötigt werden, um Daten aus einer Quelle ins BI übertragen zu können sowie die Daten zu bereinigen, zu konsolidieren und zu integrieren, um sie schließlich für Analyse, Reporting und ggf. zur Planung zur Verfügung zu stellen. Die individuellen Anforderungen Ihrer Unternehmensprozesse werden dabei durch vielfältige Ausgestaltungsmöglichkeiten des Datenflusses unterstützt.

So können Sie beliebige Datenquellen verwenden, die Daten ins BI übertragen oder direkt auf Quelldaten zugreifen, einfache oder komplexe Bereinigungs- und Konsolidierungsverfahren einsetzen und Datenablagen entsprechend Ihrer Anforderungen in der Schichtenarchitektur definieren.

Die Metadatenbeschreibung der Quelldaten wird im BI durch DataSources abgebildet. Eine DataSource ist eine Menge von Feldern, die verwendet wird, um Daten einer betriebswirtschaftlichen Einheit aus einem Quellsystem zu extrahieren und in die Eingangsschicht des BI-Systems zu übertragen oder zum direkten Zugriff zur Verfügung zu stellen. Mit der Aktivierung der DataSource erzeugt das System in der Persistent Staging Area (PSA), der Eingangsschicht des BI eine PSA-Tabelle. Somit stellt die DataSource ein persistentes Objekt innerhalb des Datenflusses dar.

Die Daten werden mithilfe eines InfoPackages in das PSA geladen, bevor Sie im BI weiter verarbeitet werden können. Im InfoPackage werden die Selektionsparameter für die Übertragung ins PSA festgelegt. Es dient im neuen Datenfluss ausschließlich dem Laden ins PSA.

Unter Anwendung der Transformation werden die Daten im BI aus einem Quellformat in ein Zielformat überführt. Somit ermöglicht Ihnen die Transformation, die Daten zu konsolidieren, zu bereinigen und zu integrieren. Die Transformation ersetzt im Datenfluss die Übertragungs- und Fortschreibungsregeln inklusive der Transferstrukturpflege. In der Transformation werden auch die Felder einer DataSource den InfoObjects des BI zugeordnet.

InfoObjects sind die kleinsten Einheiten des BI. Sie bilden die Informationen in strukturierter Form ab, die zum Aufbau von InfoProvidern benötigt werden.

InfoProvider sind persistente Datenablagen, die in der Schichtenarchitektur des Data Warehouses verwendet werden, oder Sichten auf Daten. Sie können die Daten für Analyse, Reporting und Planung zur Verfügung stellen.

Der Datentransferprozess (DTP) wird verwendet, um die Daten innerhalb des BI unter Anwendung von Transformationen und Filtern von einem persistenten Objekt in ein anderes zu übertragen. Mögliche Quellen für die Übertragung sind DataSources und InfoProvider, mögliche Ziele sind InfoProvider.

InfoArea

Element zur Gliederung der Metaobjekte im BW-System.

Jeder InfoProvider ist einer InfoArea zugeordnet. Diese Hierarchie wird in der Data Warehousing Workbench dargestellt. Auch InfoObjects können zusätzlich zu ihrer Eigenschaft als InfoProvider verschiedenen InfoAreas zugeordnet werden.

DataSource

Definition

Eine DataSource ist eine Menge von Feldern, die dem BI die Daten zu einer betriebswirtschaftlichen Einheit zur Datenübertragung zur Verfügung stellt. Technisch gesehen umfasst die DataSource eine Menge von logisch zusammengehörigen Feldern, die in einer flachen Struktur (Extraktstruktur) bzw. für Hierarchien in mehreren flachen Strukturen zur Datenübertragung ins BI angeboten werden.

Es gibt vier Typen von DataSources:

- DataSources für Bewegungsdaten
- DataSources für Stammdaten
 - ◦ DataSources für Attribute
 - ◦ DataSources für Texte
 - ◦ DataSources für Hierarchien

DataSources liefern die Metadatenbeschreibung der Quelldaten. Sie werden für die Datenextraktion aus einem Quellsystem und Übertragung der Daten ins BI oder für den direkten Zugriff auf die Quelldaten aus dem BI heraus verwendet.

Die Daten können aus einer beliebigen Quelle in der Struktur der DataSource über ein InfoPackage ins BI geladen werden. Die Festlegung, in welches Ziel die Daten aus der DataSource fortgeschrieben werden sollen, erfolgt in der Transformation. Hier nehmen Sie ebenfalls die Zuordnung von Feldern der DataSource zu InfoObjects des Zielobjekts im BI vor.

DataSources zur Datenübertragung aus SAP-Quellsystemen werden im Quellsystem definiert; die relevanten Informationen der DataSources werden durch Replikation in das BI-System übernommen. Man spricht dann von DataSource-Replikaten im BI. DataSources zur Datenübertragung aus anderen Quellen werden direkt im definiert.

Eine einheitliche Pflegeoberfläche im BI, die DataSource-Pflege, ermöglicht die Anzeige und Bearbeitung der DataSources aller möglichen Quellsystemstypen. In der DataSource-Pflege legen Sie fest, welche Felder der DataSource die entscheidungsrelevanten Informationen zu einem Geschäftsprozess enthalten und tatsächlich übertragen werden sollen.

Mit der Aktivierung der DataSource erzeugt das System eine PSA-Tabelle in der Eingangsschicht des BI. Sie können nun bereits Daten in das PSA laden. Die Selektionsparameter für das Laden der Daten ins PSA legen Sie in einem InfoPackage fest.

In der Transformation legen Sie fest, wie die Zuordnung von Feldern der DataSource zu InfoObjects des BI aussehen soll. Die Weiterverteilung der Daten aus dem PSA in weitere Ziele erfolgt über Datentransferprozesse. Hierbei werden die Regeln, die Sie in der Transformation festgelegt haben, angewendet.

Persistent Staging Area

Die Persistent Staging Area (PSA) ist die Eingangsablage im BI für Daten aus den Quellsystemen. Die angeforderten Daten werden unverändert zum Quellsystem gespeichert. Die Speicherung der Daten erfolgt in transparenten, relationalen Datenbanktabellen des BI, in denen die Request-Daten im Format der DataSource abgelegt werden. Das Datenformat bleibt unverändert, d.h. es erfolgen keinerlei Verdichtungen oder Transformationen, wie es im Falle der InfoCubes geschieht.

Eine transparente PSA-Tabelle wird zu jeder DataSource, die aktiviert wird, angelegt. Die PSA-Tabelle hat jeweils den gleichen Aufbau wie die zugehörige DataSource. Zusätzlich ist sie gekennzeichnet durch Schlüsselfelder für die Request-ID, die Datenpaketnummer und die Datensatznummer.

InfoPackages laden die Daten aus der Quelle ins PSA. Die Weiterverarbeitung der Daten aus dem PSA erfolgt über Datentransferprozesse.

InfoObject

Definition

Betriebswirtschaftliche Auswertungsobjekte werden im BI als InfoObjects bezeichnet. Sie untergliedern sich in Merkmale (z.B. Kunde), Kennzahlen (z.B. Umsatz), Einheiten (z.B. Währung, Mengeneinheit), Zeitmerkmale (z.B. Geschäftsjahr) und technische Merkmale (z.B. Requestnummer).

Verwendung

InfoObjects sind die kleinsten Einheiten des BI. Durch sie werden die Informationen in strukturierter Form abgebildet, die zum Aufbau von InfoProvidern benötigt werden.

InfoObjects mit Attributen oder Texten können selbst auch **InfoProvider** (wenn sie in einer Query werden) sein.

Struktur

Merkmale sind Ordnungsbegriffe wie z.B. Buchungskreis, Produkt, Kundengruppe, Geschäftsjahr, Periode oder Region. Sie geben Klassifizierungsmöglichkeiten des Datenbestands vor und sind damit Bezugsobjekte für die Kennzahlen. Im InfoCube beispielsweise werden die Merkmale in den Dimensionen abgelegt, welche durch Dimensions-IDs mit den Kennzahlen, die in der Faktentabelle liegen, verknüpft sind. Die Merkmale legen die Granularität (den Feinheitsgrad) fest, in der die Kennzahlen im InfoCube geführt werden. Im Allgemeinen enthält ein InfoProvider jeweils nur eine Teilmenge der Merkmalswerte aus der Stammdatentabelle. Die Stammdaten umfassen die zulässigen Werte eines Merkmals, die sogenannten Merkmalswerte.

Die **Kennzahlen** liefern die Werte, die in einer Query ausgewertet werden sollen. Dabei handelt es sich um Mengen, Beträge oder Stückzahlen. Sie bilden in einem InfoProvider den Datenteil.

Einheiten werden benötigt, damit die Werte der Kennzahlen Aussagekraft erhalten. Kennzahlen vom Typ Betrag wird immer ein Währungsschlüssel zugeordnet und Kennzahlen vom Typ Menge erhalten eine Maßeinheit.

Zeitmerkmale sind Merkmale wie Datum, Geschäftsjahr etc.

Technische Merkmale haben nur eine organisatorische Bedeutung innerhalb des BI. Ein Beispiel dafür ist die Requestnummer im InfoCube, die beim Laden von Requests als ID gezogen wird und dabei hilft, den Request wiederzufinden.

Besonderheiten der Merkmale:

Verfügen Merkmale über Attribute, Texte oder Hierarchien, so spricht man von stammdatentragenden Merkmalen. Stammdaten sind Daten, die über einen längeren Zeitraum unverändert bleiben.

Sie enthalten Informationen, die in gleicher Weise immer wieder benötigt werden. Auf diese Stammdaten kann in allen InfoProvidern referenziert werden. Es besteht auch die Möglichkeit, Merkmale mit Referenz anzulegen. Dabei liefert das Referenzmerkmal die Attribute, Stammdaten, Texte, Hierarchien, Datentyp, Länge, Anzahl und Art der geklammerten Merkmale, Kleinbuchstaben und Konvertierungsroutine für das neu angelegte Merkmal.

Eine Hierarchie wird immer zu einem Merkmal angelegt. Dieses Merkmal ist das Basismerkmal für diese Hierarchie (Basismerkmale sind Merkmale, die nicht auf andere Merkmale referenzieren). Hierarchien bieten wie Attribute eine Strukturierung der Merkmalswerte.

InfoObjects können Bestandteil der folgenden Objekte sein:

1. Bildung eines InfoProviders:

Ein InfoProvider besteht aus einer Menge von InfoObjects.

In einem InfoCube bilden die Merkmale, Einheiten und Zeitmerkmale im Wesentlichen die Schlüsselfelder, die Kennzahlen bilden den Datenteil der Faktentabelle eines InfoCubes.

In einem DataStore-Objekt bilden Merkmale in der Regel die Schlüsselfelder, sie können aber auch im Datenteil enthalten sein, zusammen mit den Kennzahlen, den Einheiten und Zeitmerkmalen.

2. Attribute zu InfoObjects

Anzeige- bzw. Navigationsattribute verfügen soll und welche Eigenschaften diese haben sollen.

Attribute sind nur vorhanden, wenn auf der Registerkarte Stammdaten/Texte das Kennzeichen Mit Stammdaten gesetzt haben.

Anzeigeattribute liefern in der Query zusätzliche Informationen zum Merkmal. Navigationsattribute dagegen werden in der Query wie normale Merkmale behandelt und können auch eigenständig ausgewertet werden.

Attribute sind schon existierende InfoObjects, die dem neuen Merkmal logisch zugeordnet sind.

InfoObjectCatalog

Definition

Ein InfoObjectCatalog ist eine Gruppierung von InfoObjects nach anwendungsspezifischen Gesichtspunkten. Es gibt 2 Typen von InfoObjectCatalogs: **Merkmal** und **Kennzahl**.

Verwendung

Ein InfoObjectCatalog ist einer InfoArea zugeordnet.

Er ist ein rein organisatorisches Hilfsmittel und dient nicht zu Auswertungszwecken.

Z.B. können alle InfoObjects, die für Auswertungen im Bereich Vertrieb eine Rolle spielen, in einem InfoObjectCatalog zusammengefasst werden. Hierdurch kann eine übersichtlichere Handhabung der evtl. sehr großen Zahl an InfoObjects in einem bestimmten Kontext erreicht werden.

Ein InfoObject kann auch in mehrere InfoObjectCatalogs aufgenommen werden.

Bei der Definition eines InfoProviders kann als Filter für die Vorlage ein InfoObjectCatalog ausgewählt werden.

Klammerung von Merkmalen

Verwendung

Bei einer Klammerung wird festgelegt, ob das Merkmal an andere InfoObjekte geklammert werden.

Die Klammerung von InfoObjekten wird in einigen Fällen notwendig, um das Datenmodell abzubilden. Einige InfoObjekte sind ohne eine Klammerung nicht eindeutig bestimmt, z.B. Klammerung des Kostenrechnungskreises an die Kostenart oder an die Kostenstelle.

Verwendung von Stammdaten und stammdatentragenden Merkmalen

Definition

Stammdaten sind Daten, die über einen längeren Zeitraum unverändert bleiben. Sie enthalten Informationen, die in gleicher Weise immer wieder benötigt werden. Im BI können die **Merkmale** Stammdaten tragen. Bei Stammdaten handelt es sich um Attribute, Texte oder Hierarchien.

Verfügen Merkmale über Attribute, Texte oder Hierarchien, so spricht man von **stammdatentragenden Merkmalen**.

Beim Anlegen eines Merkmals-InfoObjects besteht die Möglichkeit, dem Merkmal Attribute, Texte oder Hierarchien bzw. eine Kombination dieser Stammdaten zuzuweisen. Wenn ein Merkmal Stammdaten trägt, so können diese im BI-System in der Stammdatenpflege bearbeitet werden.

Sie können Merkmale als InfoProvider kennzeichnen, wenn es Attribute und/oder Texte hat. Dann steht das Merkmal als InfoProvider für Analyse und Reporting zur Verfügung.

Stammdatentypen: Attribute, Texte und Hierarchien

Verwendung

Im BI unterscheidet man drei Typen von Stammdaten:

1. Attribute

Attribute sind InfoObjects, die einem Merkmal logisch zu- bzw. untergeordnet sind und nicht in der Query selektiert werden können.



Sie können einer Kostenstelle die Attribute Kostenstellenverantwortlicher und Telefonnummer des Kostenstellenverantwortlichen (Merkmal als Attribut) sowie Größe der Kostenstelle in Quadratmeter(Kennzahl als Attribut) zuweisen.

2. Texte

Sie können zu Stammdaten Textbeschreibungen anlegen oder in das BI-System laden. Texte werden in einer Texttabelle abgelegt.



Dem Stammdatum Kostenstellenverantwortlicher ist in der Texttabelle der Name des Kostenstellenverantwortlichen zugeordnet.

3. Hierarchien

Eine Hierarchie bildet eine Zusammenfassung und Gliederung eines Merkmals nach individuellen Ordnungskriterien.

InfoProvider

Definition

Oberbegriff für BI-Objekte, in die Daten geladen werden oder die Sichten auf Daten darstellen. Diese Daten können in der Regel mit BEx Queries ausgewertet werden.

Verwendung

Als InfoProvider werden verschiedene Metaobjekte der Datenbasis verstanden, die aus Sicht einer Querydefinition in uniformer Weise als Datenlieferanten betrachtet werden können und über deren Daten folglich auch in uniformer Weise berichtet werden kann. Die Art der Beschaffung der Daten, Detaillierungsgrad oder "Nähe" zum Quellsystem im Datenflussdiagramm ist dabei von InfoProvider zu InfoProvider verschieden. Im BEx Query Designer stehen sie allerdings als gleichwertige Objekte zur Verfügung.

Unter den Begriff InfoProvider fallen zum einen Objekte, in denen Daten physisch vorhanden sind:

- **InfoCube**
- **DataStore-Objekt**
- **InfoObject als InfoProvider**

In diese InfoProvider werden über das Staging Daten geladen. Zum anderen zählen dazu auch Objekte, die keine physische Datenablage, sondern eine logische Sicht darstellen, wie:

- **MultiProvider**

InfoCube

Definition

Typ eines InfoProviders.

Ein InfoCube beschreibt einen (aus Sicht der Analyse) in sich geschlossenen Datenbestand z.B. eines betriebswirtschaftlichen Bereichs. Dieser Datenbestand kann mit der BEx Query ausgewertet werden.

Ein InfoCube ist eine Menge von relationalen Tabellen, die nach dem Sternschema zusammengestellt sind: eine große Faktentabelle im Zentrum und mehrere sie umgebende Dimensionstabellen.

Verwendung

InfoCubes werden aus einer oder mehreren InfoSources oder aus anderen InfoProvidern mit Daten versorgt. Sie stehen dann als InfoProvider für die Analyse und Reporting zur Verfügung.

Struktur

In einem InfoCube werden die Daten physisch abgelegt. Er besteht aus einer Menge von InfoObjects, die über das Staging mit Daten gefüllt werden. Seine Struktur ist die eines Sternschemas.

Integration

Auf die in einem InfoCube definierten Merkmale und Kennzahlen greifen Sie bei der [Querydefinition](#) im BEx Query Designer zurück.

Sternschema

Struktur

InfoCubes setzen sich aus einer Menge von InfoObjects zusammen. Alle InfoObjects (d.h. Merkmale sowie Kennzahlen) sind InfoCube-übergreifend verfügbar. Die Merkmale zeigen auf die Stammdaten mit ihren Attributen und Texten.

Ein InfoCube besteht aus mehreren InfoObjects und ist nach dem Sternschema aufgebaut, d.h. es gibt eine (große) Faktentabelle, die die Kennzahlen des InfoCube enthält, sowie mehrere sie umgebende (kleinere) Dimensionstabellen, in denen die Merkmale des InfoCube abgelegt sind.

Im Gegensatz zum DataStore-Objekt, das im Datenteil auch Merkmale enthalten kann, stehen im Datenteil der Faktentabelle eines InfoCubes ausschließlich Kennzahlen. Die Merkmale des InfoCubes stehen in den Dimensionen desselben.

Faktentabelle und **Dimensionen** sind miteinander verknüpft über identifizierende abstrakte Nummern (Dimensions-IDs), die im Schlüsselteil der jeweiligen Datenbanktabelle stehen. Hierdurch werden die Kennzahlen des InfoCubes auf die Merkmale der Dimension bezogen. Die Merkmale legen die Granularität (den Feinheitsgrad) fest, in der die Kennzahlen im InfoCube geführt werden.

In einer Dimension sollten Merkmale zusammengefasst sein, die sachlogisch zueinander gehören (Bezirk und Gebiet gehören beispielsweise zur regionalen Dimension). Durch Einhaltung dieses Designkriteriums wird erreicht, dass die Dimensionen weitgehend unabhängig zueinander sind und die Dimensionstabellen vom Datenvolumen her klein bleiben, was aus Performance-Gründen wünschenswert ist. Diese Struktur des InfoCubes ist für die Datenanalyse optimiert.

Faktentabelle und Dimensionstabellen sind jeweils relationale Datenbanktabellen.

Die Merkmale zeigen auf die Stammdaten mit ihren Attributen und Textbeschreibungen. Alle InfoObjects (d.h. Merkmale mit ihren Stammdaten sowie Kennzahlen) sind InfoCube-übergreifend verfügbar, im Gegensatz zu den Dimensionen, die die spezielle Organisationsform der Merkmale in einem InfoCube repräsentieren.

Dimension

Definition

Gruppierung inhaltlich zusammengehöriger Ordnungsbegriffe (Merkmale) unter einem gemeinsamen Oberbegriff. Enthält die Dimension ein Merkmal, dessen Wert die Werte aller anderen Merkmale vom betriebswirtschaftlichen Standpunkt bereits eindeutig bestimmt, so wird die Dimension nach diesem Merkmal benannt.

Verwendung

Bei der Definition eines InfoCubes werden Merkmale zu Dimensionen zusammengefasst, um sie in einer Tabelle des Sternschemas (Dimensionstabelle) abzulegen. Hierbei kann die oben genannte betriebswirtschaftliche Gruppierung Grundlage sein. Die Dimensionen sind mit Hilfe einer einfachen Fremdschlüsselbeziehung mit einem der Schlüsselfelder der Faktentabelle verknüpft.

Siehe dazu auch die Erläuterung zum **Sternschema**.

Wenn Sie einen InfoCube anlegen, dann sind per Voreinstellung die Dimensionen Datenpaket, Zeit und Einheit schon vorhanden. Die Dimension Datenpaket enthält technische Merkmale.

Zeitmerkmale und Einheiten werden automatisch den entsprechenden Dimensionen zugeordnet. Beim Aktivieren des InfoCubes werden aber nur die Dimensionen aktiviert, die InfoObjects enthalten.

Struktur

Technisch gesehen werden mehrere Merkmalswerte auf einem abstrakten Dimensionsschlüssel (DIM-ID) abgebildet, auf den sich die Werte in der Faktentabelle beziehen. Die für einen InfoCube ausgewählten Merkmale werden beim Anlegen des InfoCubes auf InfoCube-spezifische Dimensionen verteilt.

Line Item

Verwendung

Idealerweise haben Dimensionen eine– im Vergleich zur Faktentabelle – kleine Kardinalität. Jedoch gibt es Ausnahmen für diese Regel. Beispielsweise gibt es InfoCubes, in denen ein Merkmal *Beleg* verwendet wird, wobei fast jeder Eintrag in der Faktentabelle einem anderen *Beleg* zugeordnet ist. Dies führt dazu, dass die Dimension (bzw. die zugehörige Dimensionstabelle) fast genauso viele Einträge besitzt wie die Faktentabelle selbst. Man spricht in diesen Fällen auch von einer *degenerierten Dimension*.

Relationale und multidimensionale Datenbanksysteme haben im Allgemeinen Probleme, derartige Dimensionen effizient zu verarbeiten. Mit Hilfe des Kennzeichens *Line Item* lässt sich jedoch die folgende Optimierung durchführen:

Line Item: Dies bedeutet, dass die Dimension genau ein Merkmal enthält. Dies führt dazu, dass keine Dimensionstabelle angelegt wird, sondern die SID-Tabelle des Merkmals die Rolle der Dimensionstabelle übernimmt.

DataStore-Objekt

Definition

Ein DataStore-Objekt dient der Ablage von konsolidierten und bereinigten Bewegungsdaten oder Stammdaten auf Belegebene (atomarer Ebene).

Diese Daten können mittels BEx Query ausgewertet werden.

Ein DataStore-Objekt enthält Schlüsselfelder (z.B. Belegnummer, -position) sowie Datenfelder, die neben Kennzahlen auch Charakterfelder (z.B. Auftragsstatus, Kunde) enthalten können. Die Daten eines DataStore-Objektes können per Delta-Fortschreibung in InfoCubes (Standard DataStore-Objekt) und/oder in weitere DataStore-Objekte oder Stammdatentabellen (Attribute oder Texte) im selben System oder systemübergreifend fortgeschrieben werden.

Im Gegensatz zur mehrdimensionalen Datenablage bei InfoCubes werden die Daten in DataStore-Objekten in transparenten, flachen Datenbanktabellen abgelegt. Fakten- bzw. Dimensionstabellen werden nicht angelegt.

Verwendung

DataStore-Objekt-Typen im Überblick:

Typ	Struktur
Standard DataStore-Objekt	Besteht aus drei Tabellen: Aktivierungs-Queue, Tabelle der aktiven Daten, Change Log
Schreiboptimiertes DataStore-Objekt	Besteht nur aus der Tabelle der aktiven Daten
DataStore-Objekt für direktes Schreiben	Besteht nur aus der Tabelle der aktiven Daten

Standard DataStore-Objekt

Verwendung

Das Standard DataStore-Objekt wird über den Extraktions- und Ladeprozess im BI-System mit Daten gefüllt.

Struktur

Ein Standard DataStore-Objekt wird auf der Datenbank durch drei transparente Tabellen repräsentiert:

Aktivierungs-Queue: dient zur Speicherung fortzuschreibender Datensätze im DataStore-Objekt, die noch nicht aktiviert worden sind. Nach der Aktivierung werden diese Daten gelöscht, sofern alle Requests der Aktivierungsqueue erfolgreich aktiviert worden sind. Siehe auch: [Beispiel für die Aktivierung und das Fortschreiben von Daten](#).

Aktive Daten: eine Tabelle mit den aktiven Daten (A-Tabelle)

Change Log: enthält die Änderungshistorie für die Delta-Fortschreibung aus dem DataStore-Objekt in andere Datenziele, z.B. DataStore-Objekte oder InfoCubes

Die Daten gelangen über die Aktivierungs-Queue ins Change Log und werden beim Aktivieren in die Tabelle der aktiven Daten geschrieben. Während der Aktivierung werden die Requests nach ihrem logischem Schlüssel sortiert. Dadurch wird sichergestellt, dass die Daten in der richtigen Reihenfolge der Requests in die Tabelle der aktiven Daten verbucht werden.

Die Tabelle mit den aktiven Daten ist entsprechend der DataStore-Objekt-Definition aufgebaut, d.h. Schlüsselfelder und Datenfelder werden bei der Definition des DataStore-Objektes spezifiziert.

Schreiboptimiertes DataStore-Objekt

Definition

DataStore-Objekt, das nur aus einer Tabelle der aktiven Daten besteht. Die Daten werden über den Datentransferprozess geladen.

Verwendung

Daten, die in schreiboptimierte DataStore-Objekte geladen werden, stehen sofort für die Weiterverarbeitung zur Verfügung.

Struktur

Da das schreiboptimierte DataStore-Objekt nur aus der Tabelle der aktiven Daten besteht, entfällt die beim Standard-DataStore-Objekt nötige Aktivierung der Daten. Die Daten werden dadurch schneller verarbeitet.

InfoObject als InfoProvider

Definition

Ein InfoObject vom Typ Merkmal können Sie als InfoProvider kennzeichnen, wenn es Attribute besitzt. Dazu müssen Sie in der InfoObject-Pflege auf der Registerkarte Stammdaten/Texte das Kennzeichen mit Stammdaten setzen.

Die Daten werden dann mit Hilfe von Transformationsregeln in die Stammdatentabellen geladen.

Verwendung

Sie können für das Merkmal Transformationsregeln definieren und über diese Attribute und Texte laden. Es ist noch nicht möglich, Hierarchien mit Hilfe von Transformationsregeln zu laden.

Dann können Sie auch Queries für das Merkmal (genauer: für die Stammdaten des Merkmals) definieren und so über die Stammdaten reporten.

Integration

Um ein Merkmal zu einem InfoProvider zu machen, müssen Sie ihm eine InfoArea zuordnen. Das Merkmal erscheint daraufhin im InfoProvider-Baum der Data Warehousing Workbench.

MultiProvider

Definition

Ein MultiProvider ist ein Typ eines InfoProviders, der Daten aus mehreren InfoProvidern zusammenführt und sie gemeinsam für die Datenanalyse zur Verfügung stellt. Der MultiProvider enthält selbst keine Daten; seine Daten ergeben sich ausschließlich aus den zugrundeliegenden InfoProvidern, die per Union-Operation zusammengefasst werden.

Struktur

Ein MultiProvider kann sich aus verschiedenen Kombinationen der folgenden InfoProvider zusammensetzen: InfoCube, DataStore-Objekt, InfoObject, InfoSet, VirtualProvider und Aggregationsebene.

Die Daten dieser Objekte werden im MultiProvider per Union-Operation zusammengefasst, wodurch die Vereinigungsmenge der beteiligten Datenbestände gebildet wird. Es werden also alle Werte dieser Datenbestände zusammengeführt.

In einem MultiProvider muss jedes Merkmal des MultiProviders in jedem beteiligten InfoProvider genau einem Merkmal bzw. Navigationsattribut entsprechen (sofern es in diesem vorhanden ist). In mehrdeutigen Fällen muss bei der Definition des MultiProviders angegeben werden, welchem InfoObject das Merkmal des MultiProviders zuzuordnen ist.

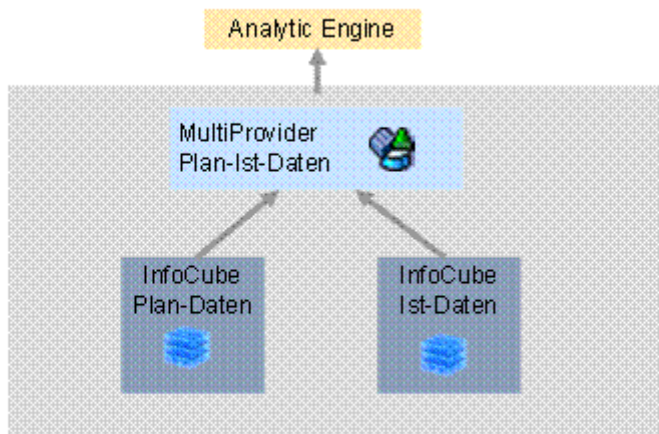
Integration

MultiProvider existieren nur als eine logische Definition. Die Daten sind weiterhin in den zugrundeliegenden InfoProvidern gespeichert.

Eine Query auf einem MultiProvider wird intern in Teil-Queries zerlegt. Dabei entsteht eine Teil-Query pro am MultiProvider beteiligtem InfoProvider. Die entstehenden Teil-Queries werden in der Regel parallel verarbeitet.

Beispiel MultiProvider: Plan-Ist-Daten

Sie haben einen InfoProvider mit Ist-Daten zu einem logisch geschlossenen betriebswirtschaftlichen Bereich sowie einen entsprechenden InfoProvider mit Plan-Daten. Sie kombinieren die beiden InfoProvider zu einem MultiProvider, um Ist- und Plan-Daten in einer Query vergleichen zu können.



Dabei handelt es sich um ein homogenes Datenmodell. Homogene MultiProvider bestehen aus technisch gleichartigen InfoProvidern, z.B. InfoCubes mit genau denselben Merkmalen und ähnlichen Kennzahlen. In diesem Fall enthält der InfoCube mit den Plan-Daten die Kennzahl Geplante Kosten und er der InfoCube mit den Ist-Daten die Kennzahl Tatsächliche Kosten.

Homogene MultiProvider stellen eine Möglichkeit der Partitionierung auf der Modellierungsebene der InfoProvider dar.

Datenextraktion aus SAP Quellsystemen

Einsatzmöglichkeiten

Zu den Datenbereitstellungsmechanismen im SAP Quellsystem gehören Extraktoren. Ein Extraktor kann die Extraktstruktur einer DataSource mit den Daten aus Datenbeständen des SAP Quellsystems füllen.

Durch Replikation wird die DataSource mit ihren relevanten Eigenschaften im BI bekannt gemacht.

Für die Datenübertragung in die Eingangsschicht des BI, die Persistent Staging Area (PSA), definieren Sie im Scheduler den Ladeprozess mit einem InfoPackage. Der Datenladeprozess wird durch ein Anforderungs-IDoc an das Quellsystem ausgelöst, wenn das InfoPackage ausgeführt wird. Wir empfehlen dabei die Ausführung über Prozessketten.

Ablauf

Es gibt anwendungsspezifische Extraktoren, die für jeweils eine mit dem BI Content des BI ausgelieferte DataSource fest programmiert sind und die Extraktstruktur dieser DataSource füllen.

Daneben gibt es generische Extraktoren, mit denen zusätzliche Daten aus dem SAP Quellsystem extrahiert und an das BI übertragen werden können. Ein generischer Extraktor weiß dabei erst zum Zeitpunkt, zu dem er aufgerufen wird durch den Namen der DataSource, zu der Daten extrahiert werden sollen, aus welchen Tabellen er die Daten in welche Struktur lesen soll. Er kann somit verschiedene Extraktstrukturen und DataSources füllen.

Unabhängig von einer bestimmten Anwendung können außerdem Stammdatenattribute, -texte oder Bewegungsdaten aus beliebigen transparenten Tabellen extrahiert werden.

Übertragung von Daten aus flachen Dateien

Einsatzmöglichkeiten

BI unterstützt die Übertragung von Daten aus flachen Dateien, Dateien im ASCII-Format (**A**merican **S**tandard **C**ode for **I**nformation **I**nterchange) oder CSV-Format (**C**omma**S**eperated **V**alue). Die Daten der flachen Datei können von einer Workstation oder von einem Applikationsserver ins BI übertragen werden.

Ablauf

1. Definition eines Datei-Quellsystem.
2. Anlegen einer DataSource im BI und definieren einer Datei die Metadaten im BI enthält.
3. Anlegen eines InfoPackages, das die Parameter für die Datenübertragung ins PSA enthält.

Transformation

Verwendung

Die Transformation ermöglicht es Ihnen, Daten zu konsolidieren, zu bereinigen und zu integrieren. Sie können Daten aus heterogenen Quellen semantisch synchronisieren.

Die Daten durchlaufen beim Laden aus einem BI-Objekt in ein weiteres BI-Objekt eine Transformation. Eine Transformation konvertiert die Felder einer Quelle in das Format des Ziels.

Funktionsumfang

Eine Transformation wird zwischen einer Quelle und einem Ziel angelegt. Als Quelle stehen Ihnen die BI-Objekte DataSource, InfoSource, DataStore-Objekt, InfoCube, InfoObject und InfoSet zur Verfügung. Als Ziel stehen Ihnen die BI-Objekte InfoSource, InfoObject, DataStore-Objekt und InfoCube zur Verfügung.

Eine Transformation besteht aus mindestens einer Transformationsregel. Verschiedene Regeltypen, Transformationsarten und Routinenarten stehen zur Verfügung, um einfache bis hochkomplexe Transformationen zu erstellen:

- Transformationsregeln: Die Transformationsregeln bilden beliebig viele Felder der Quelle auf mindestens ein Feld des Ziels ab. Dabei können verschiedene Regeltypen verwendet werden.

- **Regeltyp:** Ein Regeltyp ist eine bestimmte Operation, die durch eine Transformationsregel auf die beteiligten Felder angewandt wird.
- **Routine:** Mit Hilfe von Routinen können Sie selbst komplexe Transformationsregeln implementieren. Routinen stehen Ihnen als Regeltyp zur Verfügung. Daneben gibt es aber auch Routinenarten, mit denen Sie zusätzliche Transformationen implementieren können.

Regeltypen in der Transformation

Verwendung

Der Regeltyp bestimmt, ob und wie ein Merkmal/eine Kennzahl bzw. ein Datenfeld/ein Schlüsselfeld in das Ziel fortgeschrieben wird.

Funktionsumfang

Es stehen folgende Möglichkeiten zur Verfügung:

Direkte Zuweisung:

Das Feld wird direkt aus dem gewählten Quell-InfoObject gefüllt.

Konstante:

Das Feld wird nicht durch ein InfoObject, sondern direkt mit dem eingegebenen Wert gefüllt.

Formel:

Die Fortschreibung des InfoObjects erfolgt durch einen mit einer Formel ermittelten Wert.

Stammdaten nachlesen:

Die Fortschreibung des InfoObjects erfolgt durch Nachlesen in der Stammdatentabelle eines Merkmals, das in der Quelle mit einem Schlüssel und einem Wert enthalten ist und das entsprechende InfoObject als Attribut enthält. Über den Schlüssel werden die Attribute und ihre Werte ermittelt, diese werden dann zurückgegeben.

Routine:

Das Feld wird durch eine von Ihnen geschriebene Transformationsroutine gefüllt.

Initial:

Das Feld wird nicht gefüllt. Es bleibt leer.

Query Design: BEx Query Designer

Verwendung

Sie analysieren den Datenbestand des BI-Systems, indem Sie mit Hilfe des BEx Query Designer Queries zu InfoProvidern definieren. Durch die Auswahl und Kombination von InfoObjects (Merkmale und Kennzahlen) bzw. von wieder verwendbaren Query-Elementen (z. B. Strukturen) in einer Query bestimmen Sie, auf welche Art und Weise die Daten des gewählten InfoProvider ausgewertet werden können.

Funktionsumfang

Der BEx Query Designer umfasst folgende Funktionen:

- Sie können die Queries, die Sie im BEx Query Designer definieren, sowohl für das OLAP-Reporting als auch für das Enterprise Reporting verwenden.
- Sie können die Queries parametrisieren, indem Sie Variablen für Merkmalswerte, Hierarchien, Hierarchieknoten, Texte oder Formeln verwenden.
- Sie können die Auswahl der InfoObjects präzisieren, indem Sie
 - Merkmale auf Merkmalswerte, Merkmalswertintervalle und Hierarchieknoten einschränken
 - Formeln definieren
 - Selektionen definieren
 - berechnete und eingeschränkte Kennzahlen zur Wiederverwendung definieren
 - Lokale oder wiederverwendbare Strukturen verwenden
 - Exceptions definieren
 - Bedingungen definieren

Die wesentlichen Bestandteile der Querydefinition sind Filter und Navigation:

- Die Selektionen im Filter wirken einschränkend auf die gesamte Query. Bei der Filterdefinition wählen Sie Merkmalswerte eines oder mehrerer Merkmale oder eine Kennzahl aus. Sämtliche InfoProvider-Daten werden über die Filterselektion der Query aggregiert.
- Für die Navigation wählen Sie freie Merkmale und bestimmen den Inhalt der Zeilen und Spalten der Query. Sie legen durch diese Auswahl fest, über welche Datenbereiche des InfoProvider Sie

navigieren möchten. Die Anordnung der Zeilen- und Spalteninhalte bestimmt die Startansicht der Query.

Eine Query wird im Portal oder im BEx Analyzer nach dem Einfügen in eine Arbeitsmappe in der vordefinierten Startansicht angezeigt. Durch die Navigation der Query können Sie verschiedene Ansichten auf die InfoProvider-Daten erzeugen, indem Sie beispielsweise eines der freien Merkmale in die Zeilen oder Spalten der Query ziehen oder ein Merkmal auf einen einzelnen Merkmalswert filtern.

Durch die Definition einer Query können die Daten des InfoProvider gezielt und schnell ausgewertet werden. Je detaillierter die Query definiert ist, desto schneller werden deren Ausführung und Navigation.

Analyse & Reporting: BEx Analyzer

Verwendung

Der BEx Analyzer ist das analytische, Reporting- und Designwerkzeug des Business Explorers, das in Microsoft Excel integriert ist. Im BEx Analyzer können Sie ausgewählte InfoProvider-Daten durch Navigieren, über das Kontextmenü oder über Drag&Drop innerhalb von Queries, die im BEx Query Designer angelegt wurden, analysieren und verwenden.

Sie können auch die Oberflächen für Ihre Queries gestalten, indem Sie Design Items (Controls), wie z.B. Analysetabellen, Dropdown-Boxen und Buttons in Ihre Excel-Arbeitsmappe einfügen.

Funktionsumfang

Die Funktionen des BEx-Analyzers sind auf zwei verschiedene Modi mit jeweils einer eigenen Symbolleiste und einem eigenen Menüpfad verteilt.

- Analysemodus – Zum Ausführen von OLAP-Analysen auf Queries
- Designmodus – Zur Gestaltung der Oberfläche für Query-Anwendungen

Online Analytical Processing

Der OLAP-Prozessor im BI stellt die nötigen Funktionen und Services für eine komplexe Analyse multidimensionaler Daten sowie den Zugriff auf flache Ablagen zur Verfügung. Er bezieht die Daten aus dem Enterprise Data Warehouse und stellt diese dem BI-Frontend, dem Business Explorer, bzw. über bestimmte Schnittstellen (Open Analysis Interfaces) auch Drittanbieter-Frontends für Reporting und Analyse zur Verfügung. InfoProvider dienen dabei als Datenlieferanten. Die Abfrage auf den Daten eines InfoProviders wird durch eine Query definiert. Queries stellen somit die Grundlage der Analyse im BI dar.

Funktionen und Services

Der OLAP-Prozessor bietet Ihnen verschiedene Funktionen zur Analyse der Daten in einer Query:

- Funktionen zur Navigation in Queries wie Filter- und Aufrissmethoden (Slice and Dice), Navigation in Hierarchien (Drill-down) oder das Austauschen von Aufrisselementen (Swap)
- verschiedene Möglichkeiten zur Gestaltung der Darstellung wie die Anzeige von Ergebniszeilen oder die Gliederung in Hierarchien
- die Möglichkeit zur Formulierung von Bedingungen, um für die Analyse irrelevante Zahlen auszublenden, oder zur Festlegung von Ausnahmen (Exceptions), um kritische Werte hervorzuheben
- Durchführung von Berechnungen, wie z.B. Aggregationen, Mengen- und Währungsumrechnungen, und Verwendung von berechneten Kennzahlen oder Formeln
- Variablen zur Parametrisierung von Queries