

Neue Vergleichsstudie vom Business Application Research Center (BARC)

# Data Warehouse- wie innovativ ist SAP BW?

Teil 1 von 2



Data-Warehouse-Lösungen entwickeln sich immer stärker zu einer Basis für E-Commerce oder Customer Relationship Management (CRM). In diesen Segmenten, die eine enge Kopplung mit den operativen Enterprise-Resource-Planning(ERP)-Anwendungen erfordern, will sich auch die SAP AG mit dem SAP Business Information Warehouse (SAP BW) etablieren. Dieser Beitrag skizziert einige Erkenntnisse aus der neuen BARC-Studie über den Newcomer.

Holger Mertens, Carsten Bange und Dr. Heiko D. Schinzer

Die Veränderungen der Data-Warehouse-Produkte hören sich im Vergleich zum Megatrend der Branche, dem E-Commerce, derzeit bescheiden an. Die Religionskriege zwischen relationaler und multidimensionaler Datenbank-Technologie sind mit einem Waffenstillstand zu Ende gegangen: Fusionen oder freundliche Koexistenz sind die Folge und werden mit dem Schlagwort hybrid beworben. Doch in Details sind auch gewisse Abhängigkeiten zu den führenden Technologietrends im E-Commerce spürbar. XML gilt nun auch bei Data Warehouse als mögliches Wundermittel zur Überwindung der Metadaten-Krise und erste Reporting-Ansätze basieren inzwischen auf WAP, dem Wireless Application Protocol zur Darstellung von Internet-Seiten auf dem Handy. Was kommt nach dem Data Warehouse – neue Schlagwörter oder alte Probleme?

Derzeit bewegen sich die Anbieter sehr stark in zwei verschiedene

Richtungen. Die einen, wie Business Objects und MicroStrategy, vollziehen den Wandel vom Business-Intelligence- zum E-Intelligence-Anbieter, die anderen, zum Beispiel SAP und SAS, konzentrieren sich auf die Entwicklung von Applikationsmodulen für CRM oder Supply Chain Management (SCM) mit der notwendigen Basis einer starken analytischen Datenhaltung.

Der vom BARC durchgeführte Vergleich der zu Grunde liegenden Konzepte und Technologien ist eine große Herausforderung, da mit der Ausweitung der Modulvielfalt sowohl die Gemeinsamkeiten (vor allem auf der technologischen Ebene) als auch die Unterschiede (auf der Applikationsebene) deutlich zunehmen. Doch die Anwender in den Unternehmen werden die neue Modulvielfalt erfreut zur

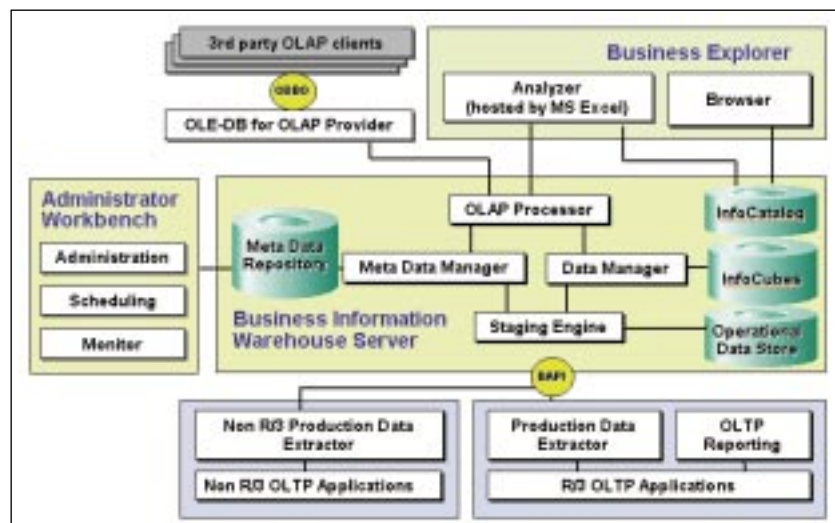


Abbildung 1: Architektur des SAP BW (Quelle: SAP AG)

Kenntnis nehmen: Besteht doch jetzt die Chance, die vor Urzeiten in Aussicht gestellten Versprechungen einer problemorientierten, aktuellen Informationsversorgung wirklich umzusetzen. Die Karawane der Anbieter-Marketiers zieht einseitigen schon weiter – Richtung Knowledge Management und Enterprise Information Portals ...

### Prüfstand für Data-Warehouse-Lösungen

Ziel der BARC-Studie „Data Warehouse“ ist der Aufbau einer aktuellen, homogenen und konsistenten Informationsbasis für Unternehmen, die sich im Rahmen der eigenen Projektierung in der Selektionsphase befinden. Dazu wurden auf Basis internationaler Marktanteilsstudien die aus Sicht des BARC führenden zwölf Anbieter (vgl. Tabelle 1) identifiziert und anhand eines einheitlichen Test-szenarios untersucht.

Im Gegensatz zu der im *is report* 01/2000 vorgestellten BARC-Untersuchung über OLAP- und Business-Intelligence-Werkzeuge stehen jetzt Produkte im Vordergrund, die mindestens einen Fokus auf der wirklichen Speicherung der Daten in einem Data Mart oder Data Warehouse haben. Die Berücksichtigung der Microsoft OLAP-Services, des Hyperion Essbase OLAP-Servers sowie der Cognos BI Suite in beiden Studien zeigt dabei aber auch ein objektives Abgrenzungsdilemma auf, dem sich auch

das BARC nicht entziehen konnte und mit der doppelten Einordnung gelöst hat. Dadurch verändert sich zwar nicht die grundsätzliche Einstellung der Autoren, dass die Analyse und nicht die Verwaltung der Daten das einzig entscheidende Ziel einer Data-Warehouse-Lösung bilden sollte. Dennoch wird bei den Beurteilungskriterien für Data-Warehouse-Lösungen ein sehr starker Fokus auf die Datenversorgung und -bereitstellung gelegt.

Auf Basis der im Kasten „Was ist ein Data Warehouse?“ skizzierten Grundarchitektur wurden knapp über 50 Bewertungskategorien abgeleitet, anhand derer die zwölf Produkte untersucht wurden. Ergebnis ist auf der einen Seite eine homogen aufgebaute, neutrale Beschreibung der einzelnen Produkte und ihrer Bestandteile sowie ein auf Basis der Untersuchungserkenntnisse beruhender Vergleich dieser Lösungen.

### Data Warehousing mit SAP BW

Nachfolgend erfolgt dabei der Versuch, auf Basis der BARC-Untersuchung das Business Information Warehouse der SAP im Vergleich zu den anderen Lösungen



Abbildung 2: Berichtsangebot im Business Explorer Browser

etwas detaillierter zu betrachten. Damit ist aber keine besonders positive Einstellung zu diesem neuen Produkt verbunden - im Gegenteil. Ziel ist es, den Unternehmen einige Hinweise zu geben, wann der Einsatz eines sehr gut zur operativen Basis (SAP R/3) passenden Data Warehouse zu präferieren ist und wann Lösungen anderer Anbieter im Vorteil sind.

Der Data-Warehouse-Ansatz der SAP weist einige Spezifika auf, die das SAP BW von allen anderen Data-Warehouse-Produkten unterscheidet. Ausgangspunkt für den Einsatz des Business Information Warehouse ist der so genannte Business Content, der aus vordefinierten Berichtsvorlagen und den zugehörigen Extraktions- und Transformationsroutinen sowie einer Datenhaltung in einem mehrdimensionalen Modell besteht. Die Berichtsvorlagen sind nach Themengebieten (zum Beispiel Ver-

**Neue BARC-Studie „Data Warehouse“**

Die BARC-Studie „Data Warehouse“ (ISBN 3-9806082-3-9) ist aktuell mit einem Umfang von 200 Seiten erschienen. Die Studie kann nur direkt bestellt werden bei:  
**Oxygen Verlag GmbH**  
 Mondstr. 2-4, D-85622 Feldkirchen  
 Tel.: +49/(0) 89/90 48 62-0  
 Fax: +49/(0) 89/90 48 62-55  
 E-Mail: info@oxygen.de  
 Der Preis: 1059,30 Mark (inkl. MwSt.; 990 Mark ohne MwSt.).

Jedes weitere Exemplar – bei gleichem Rechnungsempfänger – kostet 310,30 Mark (inkl. MwSt.; 290 Mark ohne MwSt.).

Die BARC-Studie „Data Warehouse“ vergleicht zwölf Software-Produkte zum Aufbau eines Data Warehouse.

Leseproben und Aktuelles zu den Themen Data Warehouse und Business Intelligence liefert

[www.barc.de](http://www.barc.de).



trieb oder Einkauf) oder Anwendergruppen gleich „Rollen“ (Produktmanager, Finanzbuchhalter) sortiert. Aus dem umfangreichen Katalog (zurzeit circa 450 Queries und Arbeitsmappen für 60 verschiedene Rollen) kann der Administrator des Systems dann ausgehend von den Berichtsanforderungen auswählen, welche Objekte er aktivieren will. Bei der Aktivierung werden die Definitionen für die benötigten Objekte in der Datentransportkette des SAP BW angelegt. Dies reicht von Datenhaltung und Datenobjekten bis hin zu den verschiedenen Übertragungs- und Fortschreibungsregeln. Verknüpfungen zu Datenbeständen in R/3-Systemen können über DataSources geschaffen werden, deren Metadaten in das SAP BW geladen werden. Auf

Grundlage dieser Definition wird die tatsächliche Datenbewegung dann später in einem Ladeprozess angestoßen.

Diese Vorgehensweise ermöglicht die sehr schnelle Implementierung eines Data Warehouse, da der aufwändige Prozess der Datenmodellierung und Definition des Extraktions-, Transformations- und Ladeprozesses entfällt. Die Kapitalisierung dieses Vorteils hängt allerdings entscheidend von folgenden Faktoren ab:

- Der Business Content enthält alle Objekte für die jeweiligen Berichts- und Analyseanforderungen. Der Business Content wächst ständig an: In Version 2.0A des SAP BW werden circa 2.500 InfoObjects, also Kennzahlen und Dimensionselemente be-

reitgestellt. Es besteht auch die Möglichkeit der Definition eigener Objekte, aber bei einer zu großen Zahl bietet die traditionelle Modellierung in Datenbank-Systemen größere Flexibilität.

- Es werden hauptsächlich Daten aus R/3-Systemen angebunden. Für Daten aus anderen Systemen oder nicht im Standard-Extraktor definierte R/3-Daten müssen genau wie bei anderen Produkten eigene Extraktions- und Ladeprozesse definiert werden.

Der Aufbau der integrierten Datensammlung vollzieht sich mit anderen Data-Warehouse-Projekten vollständig auf die spezifischen Bedürfnissen des jeweiligen Implementierers bezogen. Vorteil ist die Flexibilität bei Design und Aufbau, Nachteil natürlich der entstehende Aufwand. Dieser wird durch Werkzeuge wie dem Oracle Warehouse Builder oder den assistentengestützten Modellaufbau in den OLAP-Services von Microsoft allerdings gemindert.

Eine weitere Möglichkeit ist der Einsatz von Datenmodellen aus anderen Projekten. NCR, Datenbank-Spezialist für sehr große Data Warehouses, bietet so für verschiedene Branchen aus Erfahrungen abgeleitete Datenmodelle an, die häufig große Teile des gewünschten Modells bereits abbilden.

**Tabelle 1: In der Studie untersuchte Produkte**

Hersteller	Produkt	WWW-Adresse
Cognos	BI Plattform	<a href="http://www.cognos.com">www.cognos.com</a>
Hyperion Software	Hyperion Essbase	<a href="http://www.hyperion.com">www.hyperion.com</a>
IBM	Visual Warehouse	<a href="http://www.ibm.de">www.ibm.de</a>
Information Builders	SmartMart	<a href="http://www.informationbuilders.com">www.informationbuilders.com</a>
Microsoft	OLAP Services	<a href="http://www.microsoft.com">www.microsoft.com</a>
MicroStrategy	MicroStrategy	<a href="http://www.microstrategy.com">www.microstrategy.com</a>
NCR	TerraData	<a href="http://www.ncr.com">www.ncr.com</a>
Oracle	Warehouse	<a href="http://www.oracle.de">www.oracle.de</a>
Pilot	DSS Suite	<a href="http://www.pilotsw.com">www.pilotsw.com</a>
SAP	Business Information Warehouse	<a href="http://www.sap.de">www.sap.de</a>
SAS Institute	SAS System	<a href="http://www.sas.com">www.sas.com</a>
Seagate Software	Holos	<a href="http://www.seagatesoftware.com">www.seagatesoftware.com</a>

**Tabelle 2: Zertifizierte SAP-Partner zum Laden über BAPI**

Anbieter	Produkt	WWW-Adresse
Acta Technology Inc.	ActaWorks	<a href="http://www.acta.com">www.acta.com</a>
Ardent Software	DataStage load pack for SAP BW	<a href="http://www.ardent.com">www.ardent.com</a>
Evolutionary Technologies	ETI*EXTRACT	<a href="http://www.eti.com">www.eti.com</a>
Hummingbird	Genio Suite	<a href="http://www.hummingbird.com">www.hummingbird.com</a>
Informatica Corp.	PowerCenter	<a href="http://www.informatica.com">www.informatica.com</a>
Information Builders, Inc.	SNAPack Data Migrator	<a href="http://www.ibi.com">www.ibi.com</a>
Prism Solutions Inc.	Prism Connect for BW	<a href="http://www.prism.com">www.prism.com</a>
WRQ Software	Supernova for Business Warehouse	<a href="http://www.supernova.com">www.supernova.com</a>
systemfabrik GmbH	Business Warehouse Connector	<a href="http://www.systemfabrik.de">www.systemfabrik.de</a>
TSI International Software	Mercator for R/3	<a href="http://www.tsisoft.com">www.tsisoft.com</a>

**Tabelle 3: Zertifizierte Werkzeuge zum Zugriff auf SAP BW über OLE DB for OLAP**

Anbieter	Produkt	WWW-Adresse
arcplan Information Services	inSight	www.arcplan.de
Brio Technology, Inc.	Brio Enterprise	www.brio.com
Business Objects	Business Objects	www.businessobjects.com
Cognos, Inc.	Power Play SAP BW Interface	www.cognos.com
MIS Technologies	MIS Alea	www.mis.de
PointOut GmbH	PointOut for BW	www.pointout.de
Seagate Software	Seagate Info	
	Seagate Worksheet	www.seagatesoftware.com
	Seagate Crystal Reports	
	Seagate Holos	
Ton Beller AG	wwwEIS	www.tonbeller.com
Viador, Inc.	E-Portal Suite	www.viador.com

### Architektur des Business Information Warehouse

Das SAP BW adressiert wie die meisten Data-Warehouse-Werkzeuge die drei Kernprozesse Datenextraktion, Datenhaltung und -aufbereitung sowie die Darstellung und Auswertung der Daten (vgl. Abbildung 1).

Die Steuerung, Überwachung und Pflege des Extraktionsprozesses von Quelldaten wird mit der Administrator Workbench definiert und verwaltet. In dieser Workbench wird auch die Modellierung der Auswertungsobjekte vorgenommen.

Der Business-Information-Warehouse-Server sorgt im Zentrum der Architektur für die Definition und Durchführung von Transformationsschritten, beispielsweise eine Datenreinigung, und umfasst die Metadaten-Verwaltung sowie den OLAP-Prozessor zur Aufbereitung der Daten zur Darstellung in Analysewerkzeugen. Die physische Datenhaltung wird gängigen relationalen Datenbanken überlassen. Der SAP BW-Server fungiert somit als Applikationsserver und ist architektonisch vom Datenbank-Server getrennt.

Zur Darstellung und Analyse der Daten des Business Information Warehouse liefert SAP den Business Explorer in Form eines Brow-

sers zur Auswahl von Berichten sowie als Add-in in Microsoft Excel zur Berichtsdefinition und interaktiven Analyse der Daten aus.

### Datenanbindung im SAP BW

Für die Datenanbindung an das SAP BW werden drei Arten von Quellsystemen unterschieden:

- SAP R/3-Systeme,
- andere SAP BWs,
- Flat Files (ASCII) und
- Fremdsysteme (andere Datenbanken etc.).

Die Extraktion von Daten aus R/3-Systemen ab Release 3.0F geschieht über Datenextraktoren, die als Plug-ins im R/3-Applikationsserver installiert werden. Diese Extraktoren sind genau auf den Business Content, also die mitgelieferten Berichte und die dafür notwendigen Datenbestände abgestimmt.

Das ausgeprägte Know-how des Walldorfer Softwarehauses über das Wo und Wie der Speicherung von Daten im ERP-System R/3 wird so genutzt, um den gesamten Datenfluss beginnend von der Quelle im R/3-System über die Datenhaltung als InfoCubes bis zur Benutzung der Zahlen und Dimensionen in Berichten für bestimmte Adressaten zu optimieren.



### Was ist ein Data Warehouse?

Ohne im Detail den grundsätzlichen Aufbau eines Data Warehouse skizzieren zu wollen, hier in aller Kürze die wesentlichen, zu differenzierenden Elemente: Ein Data Warehouse besteht aus unterschiedlichen Komponenten, die den gesamten Prozess von der Datenanbindung und -integration über die Datenspeicherung und -verwaltung bis hin zur Datenauswertung abdecken. Die Basis für ein Data Warehouse bilden Datenquellen, aus denen Informationen extrahiert und in das Data Warehouse integriert werden. Hierbei handelt es sich in erster Linie um interne, quantitative Daten, die aus operativen Anwendungen wie zum Beispiel Produktionsplanungs- oder Vertriebssystemen stammen. Darüber hinaus bieten Anbieter von Data-Warehouse-Lösungen zunehmend die Möglichkeit, auch qualitative Informationen in Form von Dokumenten einzubinden, so dass beispielsweise zu einem überproportionalen Umsatzanstieg

bei einem Produkt Informationen über einen zurückliegenden Produkt-Relaunch angezeigt werden können. Die Überführung von Daten aus internen und externen Datenquellen in ein Data Warehouse erfolgt in drei Schritten, die als Extraktion, Transformation und Laden (ETL) bezeichnet werden. Das Datenbank-System bildet das eigentliche Data Warehouse, in dem die Daten aus unterschiedlichen Quellen gespeichert und verwaltet werden. Im Data-Warehouse-Umfeld haben sich mittlerweile zwei verschiedene Typen von Datenbank-Systemen etabliert, die zum Aufbau eines Data Warehouse geeignet sind. Hierzu gehören neben den relationalen DBMS, die häufig auch in operativen, transaktionsorientierten Anwendungen zum Einsatz kommen, so genannte multidimensionale Datenbank-Systeme. Letztere sind speziell auf die Analyse multidimensionaler Daten angepasst und erlauben daher eine performante und flexible Auswertung der Da-

ten innerhalb eines Data Warehouse. Auf Grund der Kapazitätsbeschränkungen, die sich im Bereich von einigen Gigabyte bewegen, eignen sich multidimensionale Datenbank-Systeme eher für kleinere Data-Warehouse-Lösungen (so genannte Data Marts). Data Marts enthalten üblicherweise einen eingeschränkten Datenausschnitt, der für eine Abteilung oder Benutzergruppe von besonderem Interesse ist. Die weltweite Verfügbarkeit von Informationen über einfach zu bedienende Web-Browser auf der Anwenderseite sowie die zentrale Wartung auf der Administrationsseite gehören zu den wesentlichen Vorteilen eines webbasierten Data Warehouse, was zur starken Verbreitung dieser Technologie beiträgt. Grundvoraussetzung für die Implementierung einer webbasierten Data-Warehouse-Lösung ist die Integration eines Web-Servers, der zwischen das Data Warehouse und die Client-Ebene eingefügt wird.

Andere Datenbestände können vom SAP BW selbst lediglich über Flat Files (ASCII) über eine Datei-Schnittstelle in das SAP BW überführt werden. Für die Integration weiterer Datenquellen wird eine offene Schnittstelle zur Staging Engine des SAP BW bereitgestellt: das Staging Business Application Programming Interface (BAPI). Diese Transfermöglichkeit kann sowohl von ETL-Werkzeugen als auch von eigenentwickelten Extraktionsprogrammen benutzt werden. Da die Entwicklung und Pflege von eigenentwickelten Ladeprogrammen sehr aufwändig sein kann, kommen an dieser Stelle die von SAP zertifizierten ETL-Anbieter ins Spiel (Tabelle 2). Die in der Regel nicht ganz billigen Werkzeuge unterstützen das Laden von Nicht-R/3-Daten oder R/3-Daten, die nicht von einem Standard-Extrak-

tor unterstützt werden, über die BAPI-Schnittstelle.

Im Bereich der Anbindung von Quellsystemen (inklusive SAP R/3) bieten viele andere Data-Warehouse-Werkzeuge deutlich umfangreichere Funktionalität, die teilweise bis an die Leistungsfähigkeit der oben genannten Spezialanbieter von ETL-Software heranreicht. Beispiele sind hier IBM, Oracle, Information Builders oder SAS.

Auch der Ansatz, die strukturiert vorliegende Datenbasis eines ERP-Systems für den Aufbau von vordefinierten Datenmodellen und Extraktionsroutinen zu nutzen, wird von ETL-Anbietern wie Acta oder Informatica und Data-Warehouse-Anbietern wie Cognos oder SAS nachvollzogen. Unterschied zu SAP ist allerdings, dass die Datenmodelle offen in relationalen Datenbanken abgelegt werden und so

beispielsweise auch den unterschiedlichen Anforderungen verschiedener Analysewerkzeuge gerecht werden können. Acta bietet etwa bei der Extraktion von R/3-Hierarchien (zum Beispiel Kostenstellen) drei verschiedene Extraktionsformate an, aus denen das Passende ausgewählt werden kann.

### Speicherung

Datenbasis jedes Berichtes im SAP BW sind ein oder mehrere InfoCubes, die aus einer bestimmten betriebswirtschaftlichen Perspektive die Speicherung von Kennzahlen und sie bestimmende Dimensionselemente übernehmen. Die physische Datenspeicherung erfolgt in gängigen relationalen Datenbank-Management-Systemen. Bei der Übertragung von Daten aus Vorsystemen kann angegeben werden, ob direkt in einen InfoCube

oder in den Operational Data Store (ODS) gespeichert werden soll. Der ODS ist ein Zwischenspeicher, der detaillierte Daten enthält, auf die noch Transformationsprozesse (zum Beispiel Datenreinigung, Verdichtung, Harmonisierung) angewendet werden können.

Zwischen den physisch relational gespeicherten InfoCubes und der Abfragemöglichkeit über SAP-Frontends (Business Explorer) oder die Schnittstellen (BAPI, OLE DB for OLAP) ist ein OLAP-Prozessor (ROLAP Engine) geschaltet, der für eine hohe Abfragegeschwindigkeit sorgen soll. SAP verfolgt mit dem Konzept des Applikationsservers, der eine relationale Datenbank für die Anforderungen von analytischen Abfragen optimiert, eine Vorgehensweise, die so beispielsweise auch von MicroStrategy angeboten wird. Die meisten Data-Warehouse-Werkzeuge bieten allerdings neben der relationalen auch eine physisch-multidimensionale Speicherung an, um je nach Anforderungen der Anwendung bestimmte Teile der Datenbestände in den sehr schnellen multidimensionalen Datenbanken vorhalten zu können. Beispiele sind IBM mit dem DB2 OLAP-Server, Oracle mit Express, Microsoft mit den OLAP-Services, die SAS MDDDB oder die Anbieter Pilot und Seagate.

### Präsentation und Analyse

SAP liefert mit dem SAP BW den Business Explorer als „Browser“ zur Auswahl von Berichten und als „Analyzer“ in Form eines Excel Add-ins zur Darstellung und Erzeugung von Berichten und der interaktiven Analyse von Daten (Abbildung 2).

Zusätzlich ist das Publizieren von Berichten über das Web-Reporting in Form von HTML-Seiten möglich,

und es wird ein geographisches Informationssystem (BE<sub>x</sub>-Map) angeboten. Gerade im Bereich des Internet-Zugriffs auf Data Warehouses bieten länger am Markt vertretene Data-Warehouse-Anbieter deutlich fortgeschrittenere Lösungen an. Beispiele sind Information Builders, MicroStrategy, Oracle oder SAS.

Um das SAP BW auch als Datengrundlage für andere Auswertungswerkzeuge zu öffnen, wird eine OLE-DB-for-OLAP(ODBO)-Schnittstelle angeboten. Neben Microsoft bietet inzwischen auch SAS die Unterstützung dieses Interface an. Tabelle 3 gibt einen Überblick über Auswertungswerkzeuge, mit denen ein ODBO-Zugriff realisiert werden kann.

### Beurteilung in Szenarien

Bei der Auswahl eines Data-Warehouse-Produkts steht der Entscheider vor der schwierigen Problemstellung, die Vielfalt der Produktfunktionen auf der einen Seite sowie die unterschiedlichen Anforderungen innerhalb des jeweiligen Unternehmens gegeneinander abzuwägen. Um diesen Prozess durch eine geeignete Entscheidungsgrundlage zu unterstützen, wurden im Rahmen der BARC-Studie „Data Warehouse“ insgesamt fünf Szenarien definiert, die typische Anforderungen innerhalb eines Unternehmens abbilden. Anhand des umfangreichen Kriterienkatalogs, der als Grundlage für die Beurteilung der einzelnen Produkte dient, wurden anschließend die jeweils herausragenden Produkte den entsprechenden Szenarien zugeordnet.

Die fünf Szenarien werden im nächsten *is report* erläutert. Außerdem werden zu drei Szenarien die herausragenden Produkte vorgestellt. ◀

