

# Etwas Wasser im Wein des Data Warehouses

Dr. Ulrich Himmler  
Germanenstr. 39  
65719 Hofheim

## 1 Was würde ein Aktienanalyst zu einer „Data-Warehouse-Aktie“ sagen?

Wenn man sich hie und da mal mit praktischen Dingen des Fachgebietes „Data Warehousing“ herumgeschlagen hat, ist man für Themen, die sich u. a. mit dem Zusammenspiel Data Warehouse und Information Logistik befassen, zunächst nicht so recht zugänglich. Das hängt u. U. auch mit dem eigenen beruflichen Hintergrund zusammen. Aber um das Anliegen zu verstehen, stelle man sich einmal vor, es gäbe so etwas wie eine „Data-Warehouse-Aktie“. Sie soll in diesem Gedankenspiel einen Kurs haben und auch einen Trend.

Nach meinen Eindrücken war die Aktie einmal in einem Hoch; als Analyst hätte man die Aktie wohl unter „strong buy“ eingestuft. Es verging während einer gewissen Zeit kein Tag, an dem in den relevanten und weniger relevanten Publikationen nicht ein oder mehrere Artikel zum Thema Data Warehouse erschienen wären. Der interessierte Leser war zeitweise nicht in der Lage, die Ausführungen entweder unter der Rubrik „Modeerscheinung“ oder aber „brisante Neuigkeit“ einzuordnen.

Dies allerdings ist bei relativ neuen Sachverhalten nahezu immer so.

Dann wird es langsam stiller, das Thema verschwindet nach und nach aus den Schlagzeilen bzw. wird durch andere Überschriften abgelöst. Das ist dann regelmäßig ein Zeichen dafür, dass man die Sachverhalte entweder zu den Akten gelegt hat („gestrichen Brief“), oder man ist zur intensiven sachlichen Arbeit an der Aufgabe übergegangen („bezahlt Geld“). Wenn nun - nach einer gewissen Zeit des mehr oder weniger konzeptionellen Arbeitens - Anwender<sup>1</sup> sich den „Umweg“ über das Data Warehouse sparen und „gleich“ ein Data Mart ansteuern wollen, bedeutet dies aus der Sicht der o. g. Aktie so etwas wie eine *Gewinnwarnung*, zumindest für diese Anwender ist der Stellenwert eines eigentlichen Data Warehouses gesunken. Analysten würden nicht mehr von dem ursprünglichen „strong buy“, sondern von einem „sell“ sprechen, gleichbedeutend mit einem Kurssturz.

## 2 Die Sachverhalte

Im Folgenden möchte ich die *mir* am wichtigsten erscheinenden Vorbehalte gegenüber der aktuellen Warehouse-„Szene“ *kurz* auf einen Nenner bringen, sie sollen dann meinen - zugegebenermaßen nicht vollständigen - Beobachtungen aus dem Bankensektor Deutschlands und der Schweiz gegenübergestellt werden.

- a) Man befasst sich heute schon u. a. mit der „Integration von Data-Warehouse- in Informationslogistik-Architekturen“; man betrachtet die Informationslogistik aus einer Reihe von Sichten, z. B. hinsichtlich CRM, und überall steht - im- oder explizit - ein Data Warehouse im Hintergrund. D. h., mit der Darstellung dieser heilen Welt wird ein Bild gemalt, in dem die wichtigen Informationen konsistent sind und problemlos und zeitgerecht die Entscheidungsträger erreichen. Dazu passt u. a. ein Artikel

---

<sup>1</sup> ..... oder besser „Auftraggeber oder Budgetverantwortlicher“

aus der F.A.Z. (Sonderbeilage CeBIT 2002); dort kommt man zu dem Schluss, dass erst durch die Informationslogistik die „ungeahnten Schätze“ gehoben werden können. Erst dadurch würden Daten zusammengeführt, „die zusammengehören“.<sup>2</sup>

- b) Konsequenz: Man unterstellt das Vorhandensein einer konsistenten Datenbasis. Man geht davon aus, dass die Datenquellen auch aus der anwendungsübergreifenden Sicht in Ordnung sind, man verfügt über ein ausgefeiltes Datenmanagement, die Überleitung der Daten in Richtung Data Warehouse funktioniert problemlos und ist stabil, die Ergebnisse der nachfolgenden Auswertungen - u. a. auch die angeschlossenen Data Marts - sind konsistent. Die operativen Systeme sind „nur“ deshalb nicht Gegenstand von übergreifenden Auswertungen, weil sie „nicht für derartige Auswertungen gemacht“ sind.
- c) Bevor nun die eigentlichen Probleme konkretisiert werden, sollen an einem Beispiel die Einwände transparent gemacht werden. Man stelle sich ein fiktives großes Versorgungsprojekt vor; ein Konsortium von Großfirmen möge z. B. eine Öl- und/oder Gaspipeline durch Teile Asiens und Europas bauen mit dem Ziel, die fernen Rohstoffvorkommen auf die ihnen zugedachten Märkte zu bringen. Es muss nicht weiter erläutert werden, dass es sich hier um ein in vielerlei Hinsicht herausforderndes Unterfangen handelt: Die großen Entfernungen, Grenzüberschreitungen, Gebirgsüber- und Flusssdurchquerungen wollen gut durchdacht, geplant und schließlich umgesetzt werden. Die nötigen Maßnahmen zur Aufrechterhaltung der erforderlichen Leitungsdrukke gehören auch dazu.
- d) Aber dann: Was würde passieren, wenn man nach Inbetriebnahme eines solchen Werkes feststellen müsste, dass aus den Leitungen - an deren Ende - nichts Gescheites herauskäme, jedenfalls nichts, was man für den geplanten Zweck ohne weiteres verwenden könnte? Die Abnehmer haben bestimmte Erwartungen hinsichtlich der Qualität des Outputs, von dieser Qualität abweichende Ergebnisse stellen sie vor zum Teil unlösbare Probleme: Die gesamte Investition macht ja nur Sinn, wenn die Endabnehmer des Logistiksystems mit dem schließlich gelieferten Gut das anfangen können, für das es gedacht ist. Oder man müsste die Anlage so lange einfrieren, bis durch sachdienliche Maßnahmen der „Input in die Logistik“ und damit auch der „Output“ den Erwartungen entspricht.
- e) Alle Welt weiß, dass Beispiele oft „hinken“, manchmal sogar provozieren, sie erscheinen dem einen oder anderen zu weit hergeholt etc. Dieser Vergleich hier soll niemanden verletzen oder sonst wie herabsetzen; es soll nur der Versuch gemacht werden, etwas Wasser - sozusagen homöopathisch - in den Wein des Enthusiasmus zu tröpfeln. Der Enthusiasmus soll um Gottes willen nicht geschmälert werden.
- f) Wie also kommt das störende Wasser in den Wein? Die Quellen, aus denen die Logistiksysteme ihr „Transportgut“ schöpfen sollen, sind letztlich die operativen Systeme; diese haben ihre relevanten Informationen (Daten) an ein „Hochregallager“, das Data Warehouse, abgegeben, den so genannten „single point of truth“.<sup>3</sup> Das ist jedenfalls die Theorie. Diese operativen Systeme haben es aber in sich. Sie sind aufgrund einer ganzen Reihe von Gründen vielfach nicht in der Lage, ohne weiteres das zu liefern, was sie liefern sollen. Zumindest muss man in aller Regel die Datenmenge einem Filtersystem aussetzen, um die gewünschte Qualität herzustellen. Die Einzelgründe dafür muss man sich einmal vor Augen halten, um die Stoßrichtung dieses Beitrages zu verstehen. Dabei sollte hier der Reihenfolge der aufgeführten Sachverhalte keine besondere Bedeutung beigemessen werden, sie spielt im Kontext dieser besagten Stoßrichtung keine Rolle.

### **3 Worin liegen die Probleme?**

- a) Die operativen Systeme, die Quellen der angestrebten Warehouse-Systeme, haben in aller Regel schon ein beachtliches Alter; daraus folgen - scheinbar zwingend - die „gewachsenen“ Probleme wie z. B.

---

<sup>2</sup> Alt, Guido: „Datenflut, ade“; F.A.Z vom 12-3-2002, Sonderbeilage CeBIT 2002

<sup>3</sup> Foliensatz der Fa. Merck, Datenbankrundbrief 27, GI-Fachgruppe Datenbanken

Mehrfachbelegungen von Feldern, sprechende Schlüssel, eine Vielfalt von Synonymen und Homonymen, mangelhafte Dokumentation.

- b) Die Systeme sind nahezu ohne Ausnahme als Inseln entstanden. Die Anwendungen sind nicht nur zu unterschiedlichen Zeiten entwickelt worden, die Autoren waren auch jeweils ganz andere, eine fachliche, systemübergreifende Abstimmung unterblieb weitgehend. Dabei kann man sich folgendes leicht vorstellen: Man gebe an sich gleich gut ausgebildeten Teams jeweils den Auftrag, ein operatives System zu entwickeln. Wenn nicht eine übergreifende Funktion die Abstimmung der Projekte untereinander gewährleistet, werden die Ergebnisse nahezu zwingend nichts mit dem Begriff „Integration“ zu tun haben, und dies umso mehr, je weiter die zeitlichen Startpunkte der einzelnen Projekte auseinanderfallen. Dies allerdings ist der Regelfall.
- c) Die Unterschiede in den Projektbesetzungen sowie das zeitliche Auseinanderklaffen der Arbeiten führt - auch das ist die Regel - zu einer jeweils eigenen, „spartenspezifischen“ Sprache, man formuliert eigene Definitionen der Objekte und Attribute etc. Im günstigsten Fall führt das zu Synonymen. D.h., die Autoren haben zwar von ein und demselben Attribut eine übereinstimmende semantische Vorstellung, geben ihm aber die beim jeweiligen Auftraggeber üblichen Namen. Ein Beispiel dafür sind etwa die Worte „Saldo, Kontostand, Guthaben, Restschuld etc.“. Sie bringen alle das gleiche zum Ausdruck: den wertmässigen Zustand einer Geschäftsbeziehung. Etwas anderes - und ungleich schwieriger zu behandeln - ist es, wenn die mit der Entwicklung eines Data Warehouse beauftragten Personen von einem Attribut bei gleicher Namensgebung semantisch abweichende Vorstellungen haben; es entstehen dann die Homonyme. - Die Unabhängigkeit der Entwickler führt fast zwangsläufig hier und da zu diesen Homonymen. Was z. B. ist ein Umsatz? Während man in der einen Anwendung darunter einen Buchungsposten (u. a. im Kontoauszug) versteht, verwenden die Kollegen aus dem anderen Projekt diesen Begriff in dem Sinne „der Kunde macht bei uns sehr guten Umsatz“. „Ein Unternehmen schaffte es tatsächlich, den Umsatz auf 27 unterschiedliche Arten zu definieren“.<sup>4</sup>
- d) Es gibt ein altes Lied von der fehlenden oder sehr unvollständigen Dokumentation: Oftmals weiß keiner, was ein bestimmtes Attribut, geschweige denn seine Ausprägungen bedeuten. Wenn überhaupt, muss man so etwas im Code der Anwendung ermitteln.
- e) Die Rahmenbedingungen, unter denen diese Anwendungen entstanden sind, haben vielfach zu Mehrfachbelegungen einzelner Felder geführt. Beispiel „Werbeschlüssel“: In der damaligen Landesgirokasse Stuttgart, einer Großsparkasse, hat man einen Werbeschlüssel eingeführt. Er sollte zum Ausdruck bringen, bei welchem Anlass das Adressmaterial der Kunden in das Datenmaterial der Bank gelangt ist. Zu diesem Zweck hatte jeder Schlüssel eine kurze textliche Erläuterung. Da diese Nutzungsweise des Feldes aber nicht die ganze vierstellige Kapazität benötigte, haben sich sofort andere Interessenten gefunden, die dort bestimmte - semantisch andere - Informationen ablegten. D.h., das eigentliche Feld (Attribut) hatte von da an mehrere Bedeutungen. - *Ein Eintrag jedoch bestand lediglich aus einer vierstelligen Zahl ohne die sonst übliche kurze textliche Erläuterung.* - Die Recherche war schwierig und führte schließlich zu einem Mitarbeiter, der so etwas noch wusste. Diese Menschen sind meistens Gold wert. Es stellte sich heraus, dass man mit diesem vierstelligen Begriff all jene Kunden kennzeichnete, deren Geburtsjahr - sagen wir - 1890 war. Warum? Wegen der damals üblicherweise zweistellig gespeicherten Jahreszahlen hat dann 1996 ein Stuttgarter Greis ein Schulanfängersparbuch geschenkt bekommen. Das ist zwar lustig und relativ harmlos, sogar sympatisch; es zeigt aber auch, dass man schon einigen Aufwand treiben, dass man sich den „single point of truth“ wirklich erarbeiten muss<sup>5</sup>. - Solange die einzelnen Ausprägungen eines mehrfach belegten Attributes jeweils eine textliche Erläuterung haben, führt das zwar bei der Analyse zu Mehrarbeit, ist aber noch harmlos. Ein nicht weiter erläuteter Eintrag - wie hier beschrieben - ist da schon von anderem Kaliber.

---

<sup>4</sup> Dietz, Volker: „Die zehn Gebote des Data Warehousing“, F.A.Z. vom 11-3-2002, Sonderbeilage CeBIT 2002

<sup>5</sup>U.U. hat man das Glück, dass das Data Warehouse-Projekt erst nach dem Jahrtausend-Wechsel begonnen wurde; dann ist nämlich - zumindest dieses Beispiel - aus diesem Anlass erledigt worden.

- f) Ein großes Thema sind auch die existierenden Schlüsselsysteme. Als ein Beispiel unter vielen sei hier der „zusammengesetzte Schlüssel“ eines Personenstammsatzes genannt. Er enthält u. a. die Filialnummer der Bank. Sie bringt zum Ausdruck, in welcher Filiale der Kunde sein Konto unterhält. Kommt dies bei mehreren Filialen vor, existiert der Kunde in der Personendatenbank entsprechend oft. Das wiederum hat eine Reihe von Massnahmen zur Folge, mit deren Hilfe die damit zusammenhängenden Probleme entschärft werden sollen.

Die Beispiele sind sicherlich unvollständig, auf alle Fälle sind sie *ein* Grund, weshalb ein Data Warehouse überhaupt angestrebt wird. Natürlich auch deshalb, weil die operativen Systeme nicht für Auswertungen, schon gar nicht für übergreifende Auswertungen gemacht sind. Das aber ist ein - aus meiner Sicht - technisches Problem, das hier nicht Gegenstand meiner Überlegungen ist. Mich treiben die - aus der systemübergreifenden Sicht - fachlichen Inhalte von Anwendungen um, und nur die sind für die Abnehmer von Logistiksystemen von Interesse.

## 4 Warum ist das alles eigentlich ein Problem?

Man wird mit Recht einwenden, dass man derartige Probleme eben aus Anlass der Versorgung eines Data Warehouses bereinigen muss. Wer wird schon einmal erkannte *Schwächen unbesehen auf die nächst höhere Ebene transportieren?*

In der Praxis scheint man dieser Frage keine allzu große Bedeutung beizumessen, zumindest scheint man im Hinblick auf den an sich zu treibenden Aufwand (auch an Zeit) andere Prioritäten zu setzen. Eine Reihe von Beispielen soll zeigen, dass der praktische Umgang mit dem Sachgebiet „Data Warehousing“ so seine Tücken hat.

### 4.1 Beispiel 1

Im Dialog mit Mitarbeitern eines großen deutschen Institutes wurde besprochen, dass die Bank unter „Umgehung“ des Data Warehouses ein Data Mart einführt bzw. inzwischen eingeführt hat. Man erzeugt einen Datenbestand für Auswertungen, mit deren Hilfe man auf dem Gebiet des Privatkundengeschäftes die Kunden kennen lernen möchte; es handelt sich demnach um eine Lösung, die in etwa dem „Customer Relations Management“ (CRM) dienen soll. Dabei ist es bemerkenswert, dass man für den hier genannten Zweck den vorgenannten Aufwand der Analyse der gesamten kundenorientierten operativen Systeme leisten musste bzw. geleistet hat. Man hat also bei dieser Analyse alle für ein Data Warehouse wichtigen Dinge „in der Hand gehabt“, gleichwohl hat man sich auf den Sachverhalt des angestrebten Data Marts konzentriert und beschränkt.

Man betont, dass dessen Datenbestand *keine weiteren Sichten* zulässt, z. B. sei eine Controllinganwendung ausgeschlossen. Würde man auch das abdecken wollen, hätte man es nicht nur mit weiteren (?) Daten zu tun, sondern vor allem auch mit einem weiteren Auftraggeber; und der hat - „naturgemäß“ - eine ganz eigene Sicht auf die anzustrebende Struktur der Daten (die möglicherweise Verdichtungen enthalten müssten). Wollte man diese Sicht - oder sogar noch andere Sichten wie z.B. Risc-Management - in die Aufgabenstellung einbeziehen, wäre man im Grunde wieder beim ursprünglichen Data Warehouse und seinen Terminvorstellungen, die die Auftraggeber einfach nicht akzeptieren.

In diesem Zusammenhang sind zwei weitere Sachverhalte interessant:

Die Ergebnisse der Analyse der operativen Systeme wurden nicht oder nur marginal dokumentiert; *demzufolge muss im Falle eines weiteren Data Marts der Analyseaufwand wiederholt werden.* Außerdem ist im Falle eines weiteren oder weiterer Data Marts nicht gewährleistet bzw. unwahrscheinlich, dass die dort verwendete Sprache bzw. die Interpretation der Attribute mit der aus dem ersten Data Mart übereinstimmen!

---

<sup>6</sup> Dietz, Volker: „Die zehn Gebote des Data Warehousing“, F.A.Z. vom 11-3-2002, Sonderbeilage CeBIT 2002

Es wurde kein besonderes Historisierungskonzept entwickelt. Die angestrebten Auswertungen werden sich auf den jeweils aktuellen Bestand stützen, einen so genannten „snap shot“.<sup>7</sup> Dieser wird periodisch durch einen neuen ersetzt - *nicht ergänzt*.

## 4.2 Beispiel 2

Es gibt auch Banken, die mehr als ein Data Warehouse haben. Dieser Sachverhalt hat seine Ursache in aller Regel in dem Umstand, dass bestimmte Abteilungen (Bereiche) der Institute oftmals über eine große Eigenständigkeit verfügen, auch in Dingen der IT, sie machen „ihren eigenen Kram“. Der Umstand, dass dann u. U. bestimmte operative Systeme von unterschiedlichen Data Warehouses „angezapft“ werden, hat dann die gleichen Konsequenzen wie in dem weiter oben geschilderten Sachverhalt: Es besteht die Gefahr und die Wahrscheinlichkeit, dass Attribute von unterschiedlichen Leuten auch unterschiedlich interpretiert werden, zumal dann, wenn die Dokumentation - gelinde gesagt - marginal ist. Inkonsistenzen und Unsicherheiten bezüglich der Auswertungsergebnisse sind dann natürlich die Folge.

„Für mehr als ein Data Warehouse innerhalb eines Unternehmens spricht eigentlich nichts, nur sind in einigen (großen) Unternehmen solche Inselstrukturen historisch entstanden. Da wird in einer Abteilung ein Warehouse-Projekt durchgeführt, in einer anderen Abteilung ein weiteres und eine Abstimmung gibt es nicht. Manchmal kommen auch neue Systeme hinzu, ohne dass die alten aufgegeben bzw. integriert werden. Und dann gibt es noch einzelne Warehouses für spezielle Anwendungen, wie z. B. CRM. Erst im Nachhinein merkt man dann, dass eine zentrale Lösung doch ganz sinnvoll gewesen wäre, dass Daten nicht mehr konsistent sind und dass eine Integration jetzt doch erforderlich ist.“<sup>8</sup>

„In einer konkreten Bank sind 12 Data Warehouses (!) vorhanden, die nun mit großem Aufwand nach und nach abgeglichen werden, d. h. einige werden abgeschaltet, andere werden aber vernetzt, und es gibt jetzt Richtlinien, wie Datenwürfel nun unternehmensweit zu erstellen sind. Die Heterogenität der Warehouses-Topologie bleibt jedoch erhalten.“<sup>9</sup>

In diesem Zusammenhang unternimmt man jetzt den Versuch, mittels einer in Arbeit befindlichen Software einen Datenaustausch zwischen den einzelnen Warehouses *eines* Unternehmens bewerkstelligen zu können. Es besteht der Verdacht, dass dies nur geschieht, weil man bei der Erarbeitung einzelner Warehouses bestimmte fachlich notwendige Dinge übersehen hat, die aber in „Nachbar-Warehouses“ vorhanden sind. Der Gedanke des „Time-to-Market“ fordert eben seinen Tribut. Böse Zungen sprechen von einem vorangegangenen „quick and dirty“-Vorgehen. Ob man dann wenigstens *hier* sicherstellt, dass die Semantik der abgegebenen Daten mit den tatsächlichen Anforderungen des „suchenden“ Data Warehouses übereinstimmt, müsste man mal hinterfragen.

Wenn man sich die Ergebnisse dieser - privaten - Recherchen betrachtet, trifft man immer wieder auf einen der beiden genannten Fälle. Die IT-Architekten, mit denen man redet, wissen zwar um die methodischen Schwächen eines solchen Vorgehens; aber sie sind in Zwänge eingebunden, die nun mal so etwas wie „die normative Kraft des Faktischen“ darstellen. Formulierungen wie „das machen doch alle so“ oder „wenn wir immer methodisch richtig vorgingen, werden wir nie fertig“ sind Redewendungen, die man so oder in abgewandelter Form immer wieder hört. Dabei sind die jeweiligen Gesprächspartner als kundige, gut ausgebildete und problembewusste Sachkenner zu bezeichnen. Der immer noch vorhandene Enthusiasmus steht einem zunehmenden Frust gegenüber.

In diesem Zusammenhang taucht noch ein anderer Gedanke auf:

Es gibt ein Großunternehmen aus der Automobilindustrie (BMW), das über eine große Zahl von Data Warehouses verfügt, p.a. gibt es ca. 20 *Neuprojekte*. In manchen Fällen umschleicht einen der Verdacht, dass man hie und da nur deshalb von einem Data Warehouse spricht, weil man die in diesem

---

<sup>7</sup> Möglicherweise bedeutet das Vorgehen aber auch nur einen ersten Schritt, dem die Historisierung dann irgendwann folgt

<sup>8</sup> Mail von Gunnar.Harde@offis.de

<sup>9</sup> Gunnar.Harde@offis.de

Zusammenhang fast zwingend ins Spiel kommenden Tools einsetzt. Mit anderen Worten ergibt sich folgende Frage: Wird eine Datenbank dadurch zum Warehouse, weil man sie u. a. mit OLAP auswertet?

## 5 Woran liegt das alles?

### 5.1 Time-to-Market

Datenmanager sind nicht für Geschäftspolitik zuständig, sie sollten diese - wenn irgend möglich - unterstützen. In diesem Sinne ist es nun mal so etwas wie die normative Kraft des Faktischen, wenn das Management für bestimmte Sachverhalte die Erarbeitung eines Systems anordnet, das - verglichen mit einem „richtigen“ DW - relativ schnell realisiert werden kann. Die Frage, ob man zugunsten eines Termines bestimmte Nachteile<sup>10</sup> in Kauf nehmen soll, muss letztlich die Geschäftspolitik beantworten. Ein Beispiel für derartiges Vorgehen geht aus einem Interview mit dem seinerzeitigen Projektverantwortlichen der jetzigen HypoVereinsbank in München hervor. Gefragt nach der Vorgehensweise im Rahmen der Migration der Systeme von der Hypo-Bank zur Vereinsbank lautete sein Auftrag offensichtlich dezidiert: „Termin geht vor Methode“. Es war geschäftspolitisch ganz einfach wichtig, dass sich die neue Bank nach außen möglichst bald als eine Einheit darstellen konnte, ein sicherlich einleuchtendes Motiv. D. h. aber, die dort erarbeiteten Analyseergebnisse - bezogen auf die operativen Systeme der Hypo und - vor allem - der BV wurden nicht dokumentiert und stehen für ein späteres Data-Warehouse-Projekt nicht zur Verfügung. Arbeitsgruppen, die ein „Data Warehouse“ in Angriff nehmen, machen die Arbeit noch einmal.

### 5.2 Budgetmethodik

Es liegt in der Natur der Sache, dass ein Budgetverantwortlicher und -eigentümer kein Interesse daran hat, Vorarbeiten für spätere Projekte (die nicht zu seinem Budget gehören) zu finanzieren. Im hier diskutierten Sachverhalt bedeutet das, dass die Dokumentation der Analyseergebnisse u. U. *deshalb* unterbleibt, nicht nur wegen des „Time-to-Market“-Grundes. Dabei spielt sicherlich eine Rolle, dass unter Verzicht auf „wiederverwendbare Projektergebnisse“ das Budget eher in einem „genehmigungsfähigen“ Rahmen gehalten werden kann.

Eine mögliche, politisch aber schwierig umzusetzende Lösung wäre hier die Schaffung eines übergreifenden Budgets für mehrfach nutzbare Projektergebnisse.

### 5.3 „Dezentralisierung der IT“

In einer ganzen Reihe von großen Instituten hat man den fachlichen Geschäftsbereichen jeweils „eigene“ IT-Kapazitäten zugeordnet. D. h., die ehemals zentrale IT ist jetzt zersplittert und in ihren Einzelteilen jeweils einem Geschäftsbereich nicht nur zu-, sondern auch untergeordnet; es besteht sozusagen eine „1:1“-Beziehung; der Geschäftsbereich hat seine eigene IT. Weisungsbefugnis (und Budget) hat die Leitung des Geschäftsbereichs; sie hat einerseits die IT-Kosten zu tragen, bestimmt damit aber auch, welche Projekte gemacht werden. Ob dann die in Angriff genommenen Aufgaben - mehr oder weniger gleichzeitig oder gar schon vorher - auch in anderen Bereichen angepackt werden oder wurden, spielt bei den Entscheidungen vielfach eine untergeordnete Rolle, sofern man es überhaupt weiß oder wissen will. Bei den meist vorhandenen unterschiedlichen Prioritäten, verbunden mit dem neu gewonnenen Handlungsspielraum, wird eine planerische Abstimmung vielfach als zu zeitaufwendig eingestuft. Ob dabei das mögliche „Mehrfacherfinden des Rades“ und/oder die Gefahr der Inkonsistenzen *bewusst* in Kauf genommen werden, ist offen.

---

<sup>10</sup> .... hoffentlich bewusst .....

## 5.4 Das wirtschaftliche Umfeld

Die derzeitigen Rahmenbedingungen im Finanzsektor sind zumindest teilweise gekennzeichnet durch

- A. Fusionen oder Fusionsgerichte,
- B. Margenverringering im Aktiv-/Passivgeschäft, die ehemals exorbitanten Provisionserlöse haben sich „normalisiert“,
- C. gleichzeitig erhöhte Renditeanforderungen der Aktionäre (shareholder value),
- D. Restrukturierungs- bzw. Kostensenkungsmaßnahmen.
- E. Verabschiedete oder in der Diskussion befindliche Kooperationen

Im Zweifel sind die Ursachen für das oben beschriebene Vorgehen ein Mix aus verschiedenen Motiven; es gibt eben auch hier keine monokausalen Gründe.

Aus all diesen Erscheinungen resultiert vielfach eine „gewisse“ und verständliche Zurückhaltung gegenüber an sich notwendigen *strategischen* Investitionen, die im Allgemeinen neben den finanziellen Mitteln auch eine Menge Zeit kosten. Dafür ist das planerische Umfeld durch zu viele Fragezeichen gekennzeichnet.

## 5.5 Die Kosten

Wollte man die alten operativen Systeme durch neue ersetzen, müsste man erhebliche Mittel und Zeit aufwenden, ohne dass dabei - in überschaubarem Zeitraum - ein realer betriebswirtschaftlicher Nutzen zu erzielen wäre. Die Krux ist, dass die alten Systeme teilweise schon 30 Jahre alt sind; ihre Kosten haben sich mehr oder weniger gleichmäßig auf eben diese 30 Jahre verteilt. Heute aber müsste man diese Mittel quasi sofort aufbringen bzw. zurückstellen.

Aber selbst unter Beibehaltung der alten Systeme und einer fachlich korrekten und vollständigen *Bereinigung* der operativen Daten müsste man ebenfalls große finanzielle Mittel in die Hand nehmen und den entsprechenden Zeitaufwand betreiben. Zur Erläuterung möge eine einschlägige Kostenschätzung aus dem ehemaligen Schweizerischen Bankverein dienen.<sup>11</sup>

## 5.6 Sind die Methoden auf dem Rückzug? Hat Max Vetter umsonst gelehrt?

Ab Mitte der 80-er Jahre gab es einmal eine methodische Aufbruchstimmung. Es war die Zeit, in der die relationalen Datenbanksysteme Verbreitung fanden, neue Datenstrukturen benötigt wurden und viele

---

<sup>11</sup> Schwinn, Klaus, Systor AG, Basel, auf einer periodischen „Rüdesheimer Tagung: Unternehmensmodell Banken“, Juni 1997

Anzahl Attribute \* durchschnittl. Wiederverwendung je Attribut \* durchschnittl. Anzahl Werte je Domäne \* durchschnittl. Zeitaufwand je Bewertung resp. Korrektur = notwendiger Bewertungs- und Korrekturaufwand

Dabei seien die einzelnen Faktoren wie folgt kommentiert:

Anzahl Attribute	= 20.000
Durchschnittliche Wiederverwendung	= 5
Durchschnittliche Anzahl Domänenwerte	= 10
Durchschnittlicher Zeitaufwand	= 0,5 Personenstunden

Aufwand insgesamt:

$20.000 * 5 * 10 * 0,5 \text{ Personenstunden} = 500.000 \text{ Personenstunden} = 62.500 \text{ Personentage} = 285 \text{ Personenjahre (Schweizer Werte)}$

Kosten = 285 Personenjahre \* 350.000 CHF/Personenjahr = ca. 100 Mio. CHF

Wenn man die Logik dieser Formel akzeptiert, bleibt es einem Interessenten überlassen, die einzelnen Faktorausprägungen auf sein Institut zuzuschneiden.

Informatiker so etwas wie einen konzeptionellen Neubeginn kommen sahen. Es war die Zeit, in der Max Vetter die Parole vom „Jahrhundertproblem der Datenverarbeitung“<sup>12</sup> ausgab. Er tat dies in seinem Buch, das in dieser Zeit deutlich über 10 Auflagen erfuhr; er wies dort auf die weiter oben schon angerissenen Probleme der fachlichen und sprachlichen Insellösungen hin, die Vielfalt der Homonyme und Synonyme, der Mehrfachbelegungen von Feldern etc. Und er zeigte die Wege auf, wie dies zukünftig zu bereinigen sei. Eine große Zahl von engagierten Datenmanagern machte sich an die - an anderer Stelle ungeliebte - Arbeit.

Es gibt eine Menge von Gründen, weshalb dies nicht so leicht ist. Einen an sich richtigen Neubeginn auf dem Gebiet der operativen Systeme konnte man aus Kosten- und Zeitgründen ausschließen, die geschäftspolitischen Prioritäten waren und sind andere.

Derzeit sieht es so aus, dass diese Aufbruchstimmung einer Ernüchterung gewichen ist. Der „lange Weg durch die Instanzen“ ermüdet, er frustriert auch. Und wenn dann doch einmal ein altes operatives System durch ein neues ersetzt wird, ist dies wiederum eine Insel.

Im Großen und Ganzen kann man feststellen, dass das genannte Jahrhundertproblem unverändert besteht. Und was in diesem Zusammenhang von Bedeutung ist: die Data-Warehouse-Szene ist damit konfrontiert. Es gibt Beispiele genug, dass dies nicht immer zur Kenntnis genommen wird.

## **5.7 Interessiert sich die IT für „fachlich/methodischen Unsinn“?**

Da es sich offensichtlich um ein fachliches, betriebliches Thema handelt, wird diese Frage in großen Teilen der IT (Lehre und Praxis) nicht zur Kenntnis genommen, sozusagen ausgeblendet. Man konzentriert sich auf eine so genannte „anwenderneutrale“ Sicht und arbeitet konzeptionell an der technischen Umsetzung der vielfältigen Anforderungen; dazu gehört auch eine Erweiterung dieser Sicht um z. B. eine Integration in andere, nachgelagerte Systeme, wie eben die Information Logistic. Man kann wohl sagen: je größer das Unternehmen, umso ausgeprägter ist die Arbeitsteilung, auch innerhalb der IT. Deren Vertreter reiten sozusagen sehr engagiert vor der - langsam oder vermeintlich - nachrückenden fachlichen Truppe her; diese kommt aber nicht nach.

Wo Schatten ist, ist auch Licht. Es gibt tatsächlich ein Institut, das - nach Analyse der operativen Systeme - ein Data Warehouse mit Daten versorgt, es ist der von vielen schon ungeliebte „single point of truth“. Die Hamburgische Landesbank hat dies vollbracht und sammelt derzeit erste Erfahrungen im Umgang mit dieser neuen Informationsbasis. Anschließend sollen problemorientierte Data Marts folgen. Man müsste sich einmal um die Einzelheiten kümmern. „Allerdings“: Man konzentriert sich zunächst auf einen Datenausschnitt und klammert bestimmte, hier nicht näher zu beschreibende Attribute aus.

## **5.8 Ein Ausblick**

Wenn man sich so umhört, kommt man zu dem Schluss, dass die skizzierten Schwierigkeiten weit verbreitet sind. Sie betreffen nicht nur die deutschen Banken; sie spiegeln auch die Verhältnisse in der Schweiz wider. Es stellt sich die Frage, was in dieser Sache die Zukunft bringen wird.

Wenn sich nichts ändert, ändert sich nichts; banal, aber klar. Was also müsste sich ändern?

Prognosen haben bekanntlich Schwierigkeiten mit der Zukunft. Deshalb sollen hier nur einige Stichworte genügen:

- Wie schwerwiegend muss eine Fehlentscheidung aufgrund inkonsistenter Daten sein, um ein Umdenken auszulösen?
- Werden äußere Einflüsse (z. B. Basel II) zu diesem Umdenken zwingen oder zumindest eine Diskussion in Gang bringen? In einer deutschen Landesbank ist dies derzeit der Fall.

---

<sup>12</sup> Vetter, Max: „Konzeptionelles Datendesign“, .....

- Kann ein Wandel der o. g. Rahmenbedingungen hier Abhilfe schaffen? So wie es derzeit aussieht, kann man damit zunächst nicht rechnen. Wie lange diese „Zustände“ andauern, lässt sich nicht vorhersagen. Die „Zurückhaltung“ der Entscheider scheint verständlich.
- Wird die zeit- und geldaufwändige Erneuerung der operativen Systeme möglicherweise durch das Aufkommen von Standardsoftware gelöst?
- Oder benötigt man wieder einmal einen „Missionar“ wie seinerzeit James Martin?

Die Fragen möge jeder für sich beantworten.

## Literatur

Alt, Guido: „Datenflut, ade“; F.A.Z. vom 12.03.2002, Sonderbeilage „CeBIT 2002“

Dietz, Volker: „Die zehn Gebote des Data Warehousing“; F.A.Z. vom 11.03.2002, Sonderbeilage CeBIT 2002

Harde, Gunnar: Persönliche Korrespondenz (gunnar.harde@offis.de)

Schwinn, Klaus: „Datenqualität“; Vortrag auf der periodischen „Rüdesheimer Tagung: Unternehmensmodell Banken“, Juni 1997

Vetter, Max: „Aufbau betrieblicher Informationssysteme mittels konzeptioneller Datenmodellierung“; 5. durchgesehene Auflage, Stuttgart 1989