

Erfolgskriterien für Data Warehousing- und Data Mining-Projekte in der Praxis

Data-Warehouse-Projekte sind grosse Projekte, die man schrittweise angehen muss. Welche Prioritäten würden Sie dabei setzen?

Unsere Philosophie dazu lautet «Think Strategic – Start Focused»!

Es ist unabdingbar, bereits zu Beginn die strategische Ausrichtung des Projektes festzulegen und die Unterstützung des Managements einzufordern,

GIULIANO CAMPA



Giuliano Campa ist Director Professional Services und stellvertretender Geschäftsführer.

ndern, andernfalls ist der Erfolg eines Projektes bereits vor dem Start gefährdet. Unsere Erfahrung zeigt, dass Projektzyklen 3 bis 6 Monate dauern sollten. Längere Projekte bergen die Gefahr, dass sich die ursprünglich erhobenen Anforderungen bereits wieder verändert haben. Auf Grund der vielen unbekannt Faktoren ist der erste Wurf immer der schwierigste. Wir empfehlen einen Geschäftsbereich mit nur mittlerer Komplexität zu wählen, welcher dennoch repräsentative Anforderungen mitbringt und dessen Lösung einen guten Return on Investment verspricht.

Wir unterscheiden drei Bereiche, welche den Erfolg und somit die zu setzenden Prioritäten wesentlich beeinflussen.

Zum einen sind das gewählte *Projektvorgehen* und die technischen sowie fachlichen Erfahrungen des Projektteams wichtige Schlüsselfaktoren. Die eingesetzte Methodologie (z.B. SAS Rapid Warehousing Methodology) muss das Projektteam beim Planen aller

wichtigen Aktivitäten unterstützen, sie setzt Schwerpunkte und vermeidet, dass wichtige Aufgaben zu spät oder gar nicht angegangen werden.

Zu den *Soft Factors* zählen wir vorwiegend die Kommunikation. Grossprojekte wecken erfahrungsgemäss unterschiedliche Emotionen und hohe Erwartungen. Mit einer stufengerechten Informationspolitik und einem innovativen Projekt-Marketing können die Erwartungen aller involvierten Personen proaktiv berücksichtigt werden.

Die Fachstellen sowie die Informatikabteilungen müssen als gleichberechtigte Partner von Anfang an ins Projektteam integriert werden; auf diese Art können Interessenskonflikte

tinuierliche Weiterentwicklung verliert das Warehouse-Projekt schnell an Aktualität und Akzeptanz.

Wie sieht das «Dream Team» für ein solches Projekt aus? Welche Rollen können Sie selbst besetzen, welche Rollen decken Sie mit Partnern ab und welchen Beitrag muss der Auftraggeber selbst leisten?

Je nach eingesetzter Projektmethodologie spricht man von 25 bis 30 Rollen. Je nach Projektgrösse kann eine Person durchaus mehrere Rollen wahrnehmen. Die Projektrollen teilen wir in die folgenden drei Gruppen ein:

Projektteam – Kunde	Business Projektleiter
	Executive Sponsor
	Business Unit Manager
	IT Manager
	Application Developer (Backend, Frontend, Web)
Projektteam – Service Provider	End-User-Votreter
	Warehouse Administrator
	Technischer Projektleiter
	Warehouse Architect
	Warehouse Consultant/Analyst
	Application Developer (Backend, Frontend, Web)
	Construction Manager
Projekt-Support-Team	Training Consultant
	Decision Support Specialist
	Account Manager
	Business Consultant
	Networking Specialist
	Hardware Specialist
	System Administrator
	Data Quality Specialist
	Service Provider Manager
	Data Manager
	Data Base Architect
Workshop/Project Facilitator	
Scribe	

und Konfliktpotenziale minimiert werden. Um die Erfolgchancen zusätzlich zu erhöhen, empfehlen wir, die Anwender in homogene Benutzergruppen zu segmentieren und jeweils einen Ansprechpartner zu bestimmen.

Unter *Hard Factors* verstehen wir den Einsatz von modernen Technologien wie zum Beispiel Web-Portale, Data Mining, OLAP, MIS/EIS und Ad-hoc Reporting. Attraktive und benutzerfreundliche Interaktionsschnittstellen helfen dem Anwender, sich mit der Lösung zu identifizieren und sie regelmässig zu nutzen. Sein Appetit auf funktionale und inhaltliche Erweiterung muss geweckt werden. Ohne kon-

Der Auftraggeber stellt die Rollen der Gruppe «Projektteam-Kunde». Sie stellen sicher, dass die Problemstellung und Lösungsanforderungen richtig interpretiert und die internen Rahmenbedingungen eingehalten werden.

Zusammen mit unseren Partnern können wir bei Bedarf sämtliche Service-Provider-Rollen in der geforderten Qualität und Quantität stellen. Ihr technisches Know-how und die langjährige Erfahrung garantieren eine erfolgreiche Umsetzung der jeweiligen Kundenbedürfnisse.

Das Projekt-Support-Team unterstützt das Projektteam punktuell in fachlichen und technischen Detailfra-

gen. Die Know-how-Erwartungen sind daher sehr weitreichend. Je nach Wunsch kann einen Grossteil dieser Rollen durch unsere Mitarbeiter oder unsere Partner abgedeckt werden.

Um eine Abhängigkeit des Kunden nach Projektende zu vermeiden, versuchen wir in jedem Fall ein ausgewogenes Projektteam zusammenzustellen. Im «Dream Team» sollten auch Entwickler des Auftraggebers beteiligt sein. Sie lernen während des Projektes den Umgang mit den Eigenheiten eines Warehouse-Projektes und dessen technischen Lösungen.

Der Nutzen eines Data Warehouse steht und fällt mit der Qualität der Daten, die dafür zur Verfügung stehen. Welche Punkte muss man besonders beachten bei der Datenbereitstellung?

Die Datenqualität ist in der Tat einer der wichtigsten Einflussfaktoren. Bereits während der ersten Projektphase (Assessment) muss die Qualität der Daten und deren Integrierbarkeit geprüft werden. Oft müssen parallel zum Warehouse-Projekt zusätzliche Data-Quality-Projekte lanciert werden, damit die Daten überhaupt wiederverwendet werden können. Dabei kann es sich von simplen Bereinigungen bis hin zur Optimierung der Datenerfassungsprozesse bewegen. Mit der Anzahl von Datenquellen und undokumentierten Formeln und Funktionen steigt die Wahrscheinlichkeit, dass die Datenqualitätsproblematik ein treuer Projektbegleiter sein wird.

Bei der Datenakquisition unterscheidet man zwischen Push- und Pull-Strategie. Bei der traditionellen Pull-Variante stellt das Quellsystem die Daten zur Verfügung. Hier ist es wichtig, ein Service Level Agreement (SLA) zu definieren. Das SLA muss die Verantwortlichkeiten, die zu liefernden Daten und die vereinbarten Qualitätsmerkmale festlegen.

Oft sind Kenntnisse über den Lebenslauf von Daten und ihren Systemen von Vorteil, wurden Daten und Systeme auf Grund von internen Reorganisationen, Strategiewechsels oder Firmenakquisitionen mehrfach mutiert, kann davon ausgegangen werden, dass die Datenqualität darunter gelitten hat.

Erkenntnisse, welche aus den Datenbereinigungsregeln gesammelt werden, müssen den Verursachern gemeldet werden, damit die Ursache behoben werden kann. Andernfalls drohen die zahllosen Regeln das Warehouse-

Projekt mit zusätzlichen Entwicklungsaufwänden unnötig zu verzögern und die Wartbarkeit negativ zu beeinflussen.

Die Qualitätsprüfung ist keine einmalige Angelegenheit. Auch während des produktiven Betriebes muss sichergestellt werden, dass die Daten auf ihre fachliche, wie aber auch technische Richtigkeit geprüft werden. Hierfür gibt es eine breite Palette von technischen Lösungen, trotzdem muss die Verantwortung hierfür klar geregelt sein.

Welche bankspezifischen Data-Warehouse-Anwendungen könnten Sie realisieren? Haben Sie bereits ein Datenmodell für die Bankbranche?

Grundsätzlich können wir alle Problemstellungen im Bereich von Business Intelligence lösen. Es gibt kaum Bereiche, in denen wir nicht mit unseren Lösungen echten Business Value für den Auftraggeber generiert haben. In der Schweiz haben wir, zusammen mit unseren Partnern, Anwendungen und grosse Erfahrungen in den Bereichen Total Financial Management (Controlling, Budgeting, Konsolidierung), Risk Management, Kreditmanagement, Geldwäscherei, Fraud Detection, Customer Retention u.v.m.). Unser weltweites Resource-Sharing-Programm ermöglicht uns, die jeweiligen Experten zu verpflichten.

Datenmodelle der operativen Systeme reflektieren sehr stark die Prozesse einer Unternehmung. Data-Warehousing-Modelle hingegen verfolgen ein anderes Ziel und sind wesentlich abstrakter. Unsere generischen Datenmodelle können trotzdem nur als Grundlage dienen, sie müssen an die jeweiligen Gegebenheiten angepasst werden. Zusammen mit unseren Partnern haben wir für verschiedene Bereiche Vorlagen erarbeitet. Ein Teil der effektiven Modelle sind verständlicherweise vertraulich und können nicht 1:1 wiederverwendet werden.

Mit Data Mining lassen sich in grossen Datenbeständen bisher unbekannte Zusammenhänge entdecken. Welche Methoden halten Sie dabei für die besten?

Data Mining umfasst einerseits sehr viele Anwendungsbereiche wie Marketing oder Finanzen, andererseits auch viele unterschiedliche Fragestellungen: von der Analyse charakteristischer

Kundensegmente bis zur Modellierung von Kaufwahrscheinlichkeiten oder Risikoabschätzungen. Daher hängt der Erfolg von Data-Mining-Projekten nicht von einer bestimmten Methode, sondern von der Wahl der richtigen Methode für die jeweilige Fragestellung ab. So erfordert beispielsweise die Durchführung einer Assoziationsanalyse grundsätzlich andere Verfahren als diejenigen, die für eine Segmentierung oder für ein Risiko-Scoring benötigt werden. Darüber hinaus unterscheiden sich auch Methoden, die prinzipiell für die gleichen Aufgabenstellungen eingesetzt werden können – z.B. Neuronale Netze und Entscheidungsbaumverfahren –, grundsätzlich hinsichtlich der Einfachheit der Anwendung, der Interpretation der Ergebnisse und der Modellierungsfähigkeit voneinander. Daher ist es nicht möglich, a priori eine Methode als «die Beste» zu definieren. Um das am besten geeignete Verfahren auswählen zu können, muss vielmehr die Möglichkeit gegeben sein, unterschiedliche Verfahren verwenden und deren Ergebnisse mit geeigneten Mitteln untereinander vergleichen zu können. Daher sind Faktoren wie Integration und Skalierbarkeit viel entscheidender für Erfolge mit Data Mining als eine bestimmte Methode. SAS bietet mit seiner mehrfach ausgezeichneten Lösung für Data Mining alle diese notwendigen Funktionalitäten.

Ein weiterer Aspekt, der entscheidenden Einfluss auf das Gelingen von Data-Mining-Projekten hat, ist das Vorhandensein einer geeigneten Methodologie, die optimalerweise auf ganz spezifische Fragestellungen fokussiert ist. Hierzu hat SAS so genannte «Knowledge Solutions» entwickelt, die aus einem Paket von Software, spezifischem Know-how und Dienstleistungen bestehen, die branchenspezifische Fragestellungen adressieren, wie z. B. Churn-Management, Credit Scoring oder Cross-Selling.



SAS Institute AG
Ruchstuckstr. 6
8306 Brüttsellen
Tel. 01/805 74 74
Fax 01/805 74 99
E-Mail: info@sch.sas.com
Web: www.sas.com/switzerland