

2. Workshop Open Source BI 2011 Hochschule Karlsruhe

Beispielhafte BI-Projekte im Mittelstand mit OSS
Alexander Pohl

[Index](#)



» Kurz zu mir

» Beispielhaftes Projekt eines Hotelkonzerns

- Der Kunde
- Scope und Umfang, Ist und Soll
- Risikofaktoren und Gegenmaßnahmen, Plan, Rollen
- Vorgehensmodell – was unterscheidet ein OSS BI Projekt

» Die Implementierung

- Welche Hürden wurden genommen, wie, welche Alternativen
- Argumente der gewählten Lösung
- OSS Spezifika

» Projekt Stader Saatzucht mit Spaltenbasierter DBMS





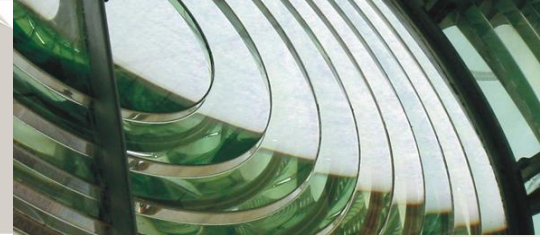
» Start bei BDO, 6 Jahre als Unternehmensberater

- Zunächst Juniorberater, nach 2 Jahren Experte für Systemsicherheit und Qualität von IT- und Softwareentwicklungsprozessen
- Im 4. Jahr Teilprojektleiter bei der Berufsgenossenschaft BGF
- Im 5. Jahr Aufbau Geschäftsbereich „Neuer Markt“

» 1997 Gründung uniware consulting/heute 3A-Strategy:

- Groupware mit Workflow, Sicherheit der Kommunikation im Internet
- PM und Krisenmanagement von IT-Projekten
- Konsolidierung von IT-Portfolios, ROI Betrachtung
- Outsourcing und Offshoring von Projekten
- Unterstützung beim Controlling und Due Diligence
- PL und Implementierung von Controlling Werkzeugen, DataMarts
- Toolkenntnisse: Infor/Alea, Microsoft, Palo, SQL, TM1 und viele Open Source Projekte





» Hotelkonzern mit 50 Standorten

- Mehrere Sprachen und Währungen
- Heterogene Systemlandschaft:
 - Verschiedene Zeiterfassungsgeräte
 - Unterschiedliche Telefonanlagen und Abrechnungssysteme
 - Mehrere Kassensysteme
- Dezentral organisiert
- Überlastete IT, sehr mittelständisch geprägt
- Kontakt über Partner für Personalzeiterfassung

» Aufbau MIS mit Data Mart für:

- Zentraler Zugriff auf Daten der operativen Systeme
- Abschottung der personenbezogenen Daten





» Scope und Umfang:

- **Start:** Webportal für die Zentrale mit Zugriff auf alle HR Daten
- **Dann:** Erweiterung um Abrechnungsdaten aus Buchhaltung
- **Zuletzt:** Integration mit Fidelio (micros)

» Ausgangssituation und Ziel:

- Keine Transparenz, unregelmäßig werden die Zahlen unter hohem Aufwand aus Plan und Ist zusammen getragen
- Trennung der Sichten – Betriebsrat, Bedienung durch Controller
- Unternehmensweites MIS bzw. Dashboard
- Fehlerwahrscheinlichkeit sehr hoch, vieles manuell, etwa KK
- Tagesgenaue Anpassungen nicht möglich, etwa Personalplanung

» Business Benefit:

- Planabweichungen frühzeitig für Gegenmaßnahmen darstellen
- Reduktion operativer Tätigkeiten – Controlling strategisch nutzen
- Controlling vor-Ort/dezentral ermöglichen



Generell gilt bei BI bzw. Data Mart Projekten - Risikofaktoren bzw. typische Stolperfallen



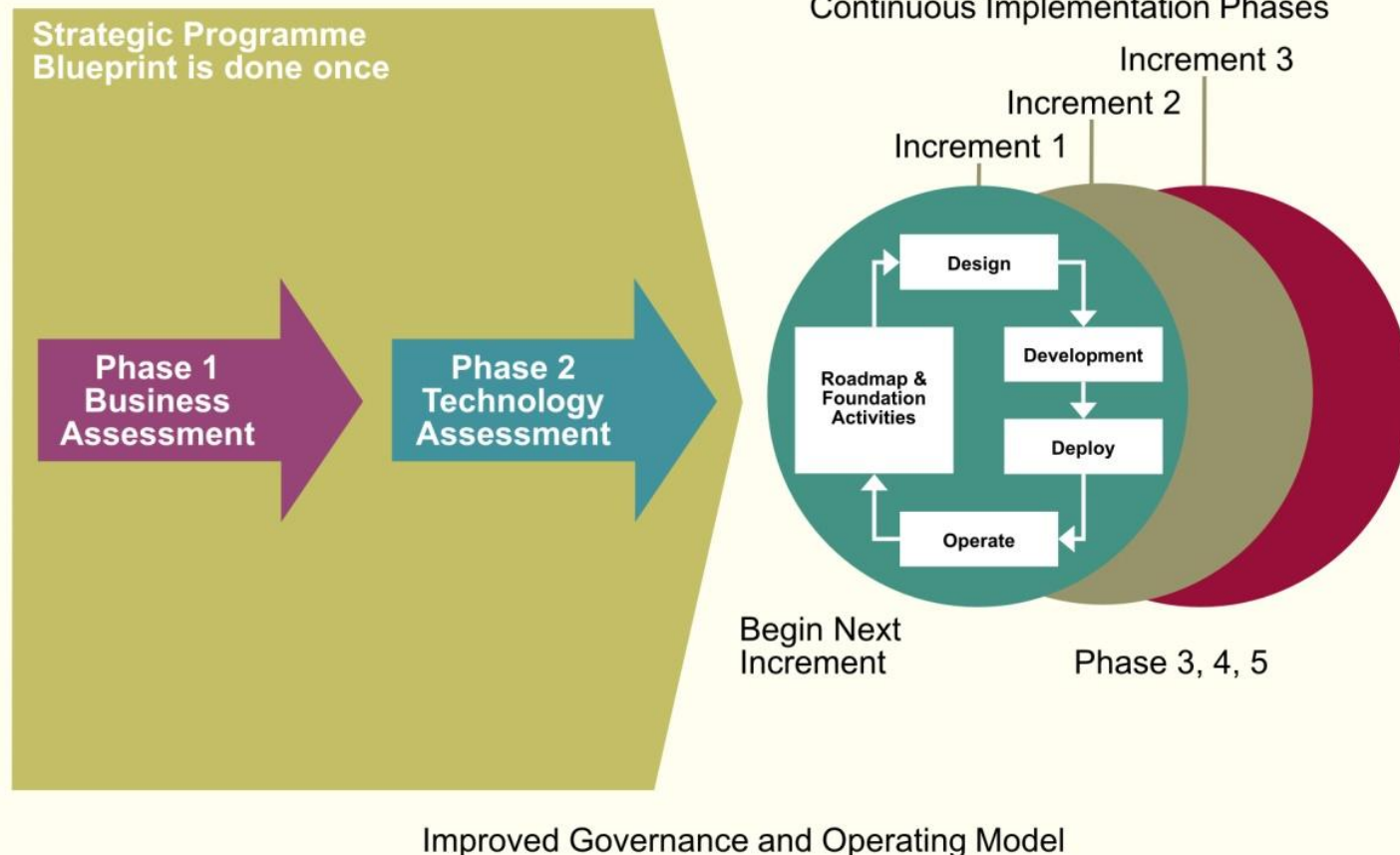
- » Teuer und Budgetüberschreitung -> AUCH im MS!
 - Zusätzliche Aufgaben mit zus. Personal & externe Beratung
 - Bindung der internen Kräfte - Know-how Träger sind Flaschenhals
 - Zusätzliche Verantwortlichkeiten bzw. neue Organisationsstrukturen
 - Fachkompetenz plus Technologie aus Soft- & Hardware
- » Notwendig: 5/7 Phasen – Loop bzw. DWH Lifecycle:
 - I. Interne Validierung bzw. Kosten-/Nutzenanalyse
 - II. Planung strategisch
 - III. Detaillierte Analyse aus Sicht des Geschäftszwecks – was sind die konkreten Ziele aus Sicht der Anwender und die davon abgeleiteten Anforderungen
 - IV. Blaupause und Planung operativ
 - V. Entwicklung des geplanten Produktes
 - VI. Test, Implementierung, Abnahme und Real World Einsatztest des Produktes – auch hinsichtlich ROI und laufender Kosten



Bei der Vorgehensweise von DWH bzw. DataMart Projekten gibt es viele erprobte Standards - wir halten uns an Bewährtes:



Information Development through the 5 Phases of MIKE2.0



Konkret auf unser Projekt bezogen:
Sehr unstetes, heterogenes Umfeld, knappe
Ressourcen gepaart mit großen Hoffnungen



» Projektrisikofaktoren:

- Technik bzw. OS unbekannt
- Einbindung verschiedenster Anbieter
- IT-Mannschaft bereits heute überlastet
- Hohe Erwartungshaltung bei engem Budget
- Ungeschulte Anwender

» Gegenmaßnahmen:

- Entwicklung eines PoC mit Beschränkung auf HR (System bekannt)
- Zeitgleich bereits Kontakt mit Anbietern herstellen, Möglichkeiten und auch deren Nutzen aufzeigen
- Bei allen Aktivitäten möglichst wenig IT-Personal nutzen, bzw. jeweils mit langem Vorlauf
- Teile der Komponenten und Transformationen per Offshore
- Permanente, stringente Überwachung, Anpassung Budget, Plan und RAEW Matrix
- Separater Aufwand vorgesehen für Schulungsunterlagen und CMS





» Proof of Concept (PoC):

- Kurze Aufnahme der gesamten Systemlandschaft
- Analyse der bereits vorhandenen Schnittstellen HR
- Modellierung DWH, Transformationen für bekanntes HR System
- Aufbau des Webportals für Plan- und Istzahlen HR

» MIS für Konzern-Controlling:

- Implementierung unserer **PDI/Kettle Schnittstelle zu Datev KaReWe**, Erstellung DWH Datenstrukturen Financial Data Mart
- Import Daten und Cleansing, Definitionen & Plausis erstellen
- Planungs- und Konsolidierungswürfel bereitstellen mit unserem **Palo/Kettle Plugin**

» MIS konzernweit:

- Analyse der notwendigen Daten und Schnittstellen zu anderen OS
- Aufwandsschätzung und Projektplan



PoC nach 2 Monaten, Corporate Controlling nach weiteren 2 Monaten, CPM unternehmensweit nach 5 Monaten



Project Phases and RAEW Matrix					Implementation		Implementation		
Function	Resources	Assessment			Change & Execution		Change & Execution		
		RAW	PoC	Blueprint	RAW	Optimizing	RAW	Optimizing	Optimizing
Strategy & Leadership	AP, CC	RAW	RAW	RAW	RAW	RAW	RAW	RAW	RAW
	IT	AE	EW	AE	EW	W	AE	EW	W
Value & Risk Management	DS	RAW	RAW	RAW	RAW	RAW	RAW	RAW	RAW
	CC	AW	A	AW	A	AW	AW	A	AW
Requirements Architecture & Design	IT	EW	EW	EW	EW	EW	EW	EW	EW
	AP	RAW	AW	RAW	AW	A	RAW	AW	AW
	DS	AW	W	AW	W	A	AW	W	AW
	ML	EW	W	EW	W		EW	W	
	CC	AW	RAW	AW	RAW		AW	RAW	
	IT	E	E	EW	E		EW	E	
Engineering	AP	RAW	RA	RA	RA	RA	RA	RA	RA
	ML	W	AE	AEW	AE	AE	AEW	AE	AE
	OF		EW		W	EW		EW	W
	CC	W	W	W	W	W	W	W	W
	IT	EW	EW	EW	EW	EW	EW	EW	EW
Documentation	AP	RAW	RAW	RAW	RAW	RAW	RAW	RAW	RAW
Training	DS	RAW	RAW	RAW	RAW	RAW	RAW	RAW	RAW

RAEW Matrix/Who:

•

• W - Work

AP - Alexander Pohl
DS - Dirk Schirmer
OF - Offshore Team
ML - Matts Lagerwall
IT - IT-Abteilung
CC - Corporate Controlling



Lösung auf der Grundlage neuer Technik, Perspektive zur unternehmensweiten Nutzung bzw. Roll-Out



» Integration diverser Daten bzw. –quellen per OSS-ETL:

- Oracle DBMS, MS SQL Server (Express)
- Spreadmarts mit Excel, Access
- Text Export Dateien aus Datev KaReWe
- Mit Pentaho PDI: Fertiges Financial DM – Anbindung Datev & Palo

» Erstellung Strukturen für DWH in MySQL und Infobright:

- Gleiches SQL Interface, geringe Lernkurve, OSS Variante für Beginn
- PostgreSQL mit Greenplum bot nicht die Small Footprint Infrastruktur geprägt von Einfachheit und Performance

» Webportal mit SpagoBI:

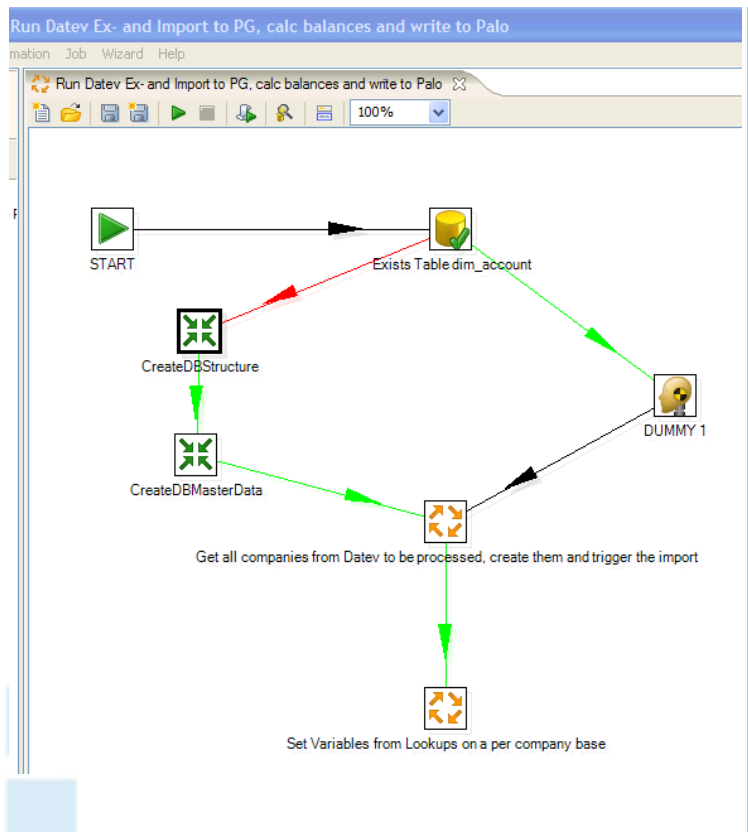
- Schlechte Doku, eingeschränkte Funktionalität
- Definition der Must's, Beschränkung im Funktionsumfang



Erstellung per PDI automatisiert unser Datev DataMart



Der zentrale Job zur Steuerung des gesamten ETL Prozesses:



Vollautomatische Anlage des Financial Data Marts in PostgreSQL in der Transformation „Create DB Structure“:

```
CREATE SEQUENCE "public"."facts_gl_pl_accounting_nation_id_seq";
CREATE SEQUENCE "public"."facts_gl_pl_accounting_costcenter_id_seq";
CREATE SEQUENCE "public"."facts_gl_pl_accounting_costunit_id_seq";
CREATE SEQUENCE "public"."facts_gl_pl_accounting_profitcenter_id_seq";
CREATE SEQUENCE "public"."dim_cc_id_seq";
CREATE SEQUENCE "public"."dim_cc_company_id_seq";
CREATE SEQUENCE "public"."dim_hier_cc_id_seq";
CREATE SEQUENCE "public"."dim_account_id_seq";
CREATE SEQUENCE "public"."dim_account_teckey_name_schema_seq";

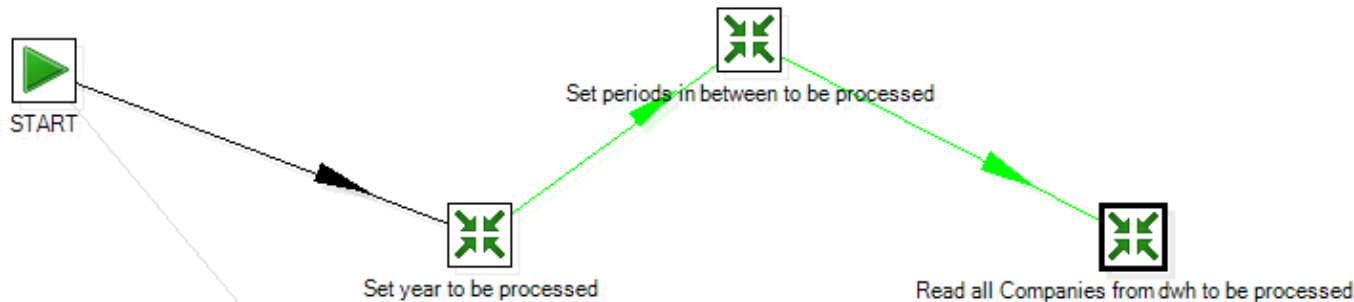
CREATE TABLE public.dim_hier_cc (
  id int4 NOT NULL DEFAULT nextval('dim_hier_cc_id_seq'::regclass),
  cc_from_id int4 NOT NULL DEFAULT nextval('dim_cc_id_seq'::regclass),
  company_id int4 NOT NULL DEFAULT nextval('dim_company_id_seq'::regclass)
);

CREATE TABLE "public"."dim_account" (
  "schema" varchar(6) NOT NULL,
  "name" varchar(11) NOT NULL,
  "name_int" integer NULL,
  "label_english" text NULL,
  "label_german" text NULL,
  "id" int8 NOT NULL DEFAULT nextval('dim_account_id_seq'::regclass),
  "teckey_name_schema" int8 NOT NULL DEFAULT nextval('dim_account_teckey_name_schema_seq'::regclass),
  "version" int4 NULL
);
```

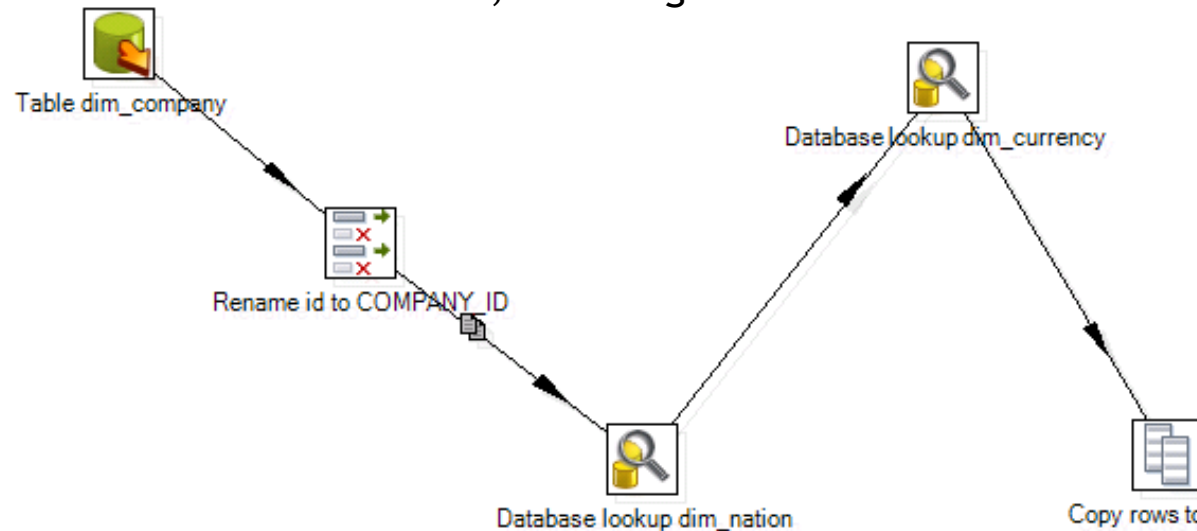
PDI: Steuerdaten festlegen, Masterdaten einlesen



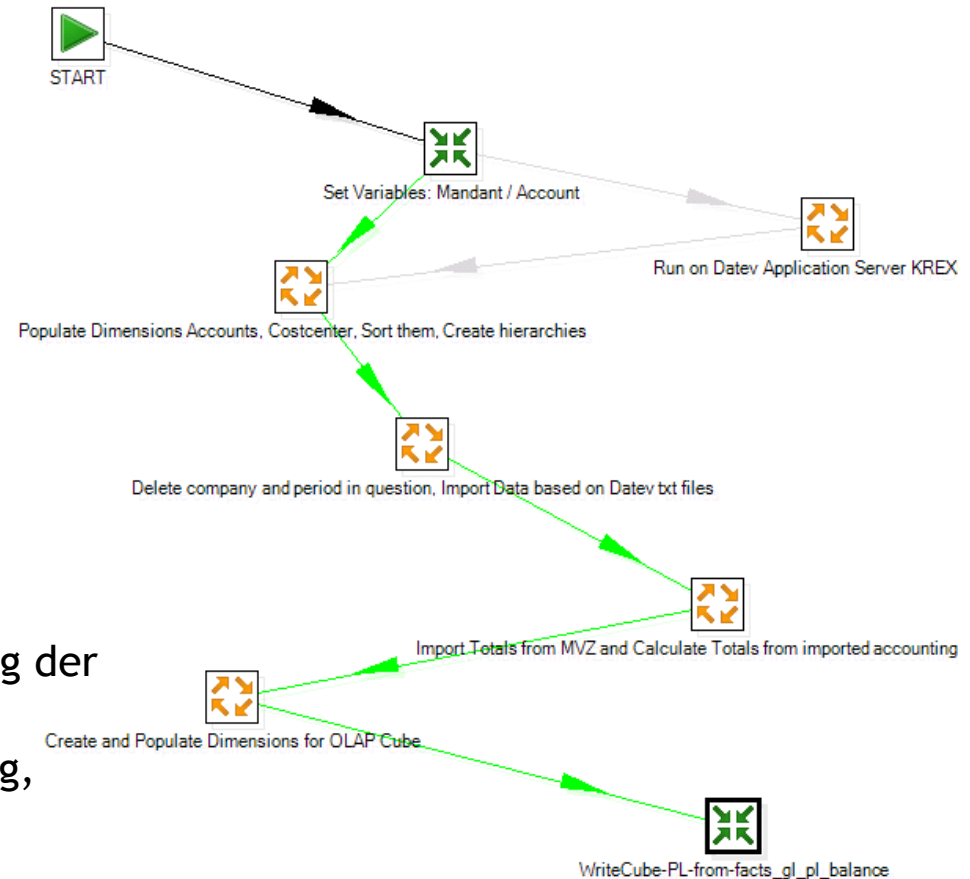
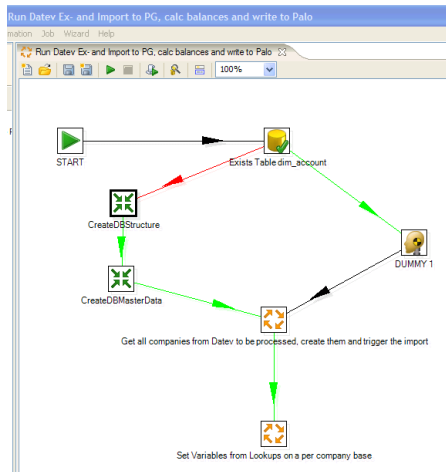
Einstellungen vornehmen - Jahr, Monat und Mandant Job „Get all Companies“:



Transformation „Read all Companies ...“ - Auswerten der Firmen/Mandanten in KaReWe, Auswerten der Steuerdaten/Masterdata hinsichtlich Land, Währung:



PDI: Umgebungsvariablen mit Steuerdaten, Aktualisierung und Import der Finanzen



Job „Set Variables ...“ - Loop pro gefundenem Mandanten mit Einstellung der notwendigen Parameter, Ergänzen der Elemente und Attribute falls notwendig, Import der Buchungssätze, Saldieren der Buchungen auf Periodenebene, Insert in Datenbank und Load in den OLAP Cube:



Zwischenergebnis Star Schema

Tabelleninhalt mit Buchungssatz im Data Mart



Zentrale Buchungstabelle (facts_GeneralLedger), Spalten für Konsolidierungsebene, Mandant, Sparte, Land, Kostenstelle (Kostenstelle kann nur aus dem einzelnen Buchungssatz ausgelesen werden, daher kann die Saldierung aus KaReWe in Form der Monatsverkehrszahlen (MVZ) nicht herangezogen werden, usf.
mit Bezug auf dim_conslevel, dim_period, dim_version:

The screenshot shows a database management tool interface. On the left, a tree view displays the following schema structure:

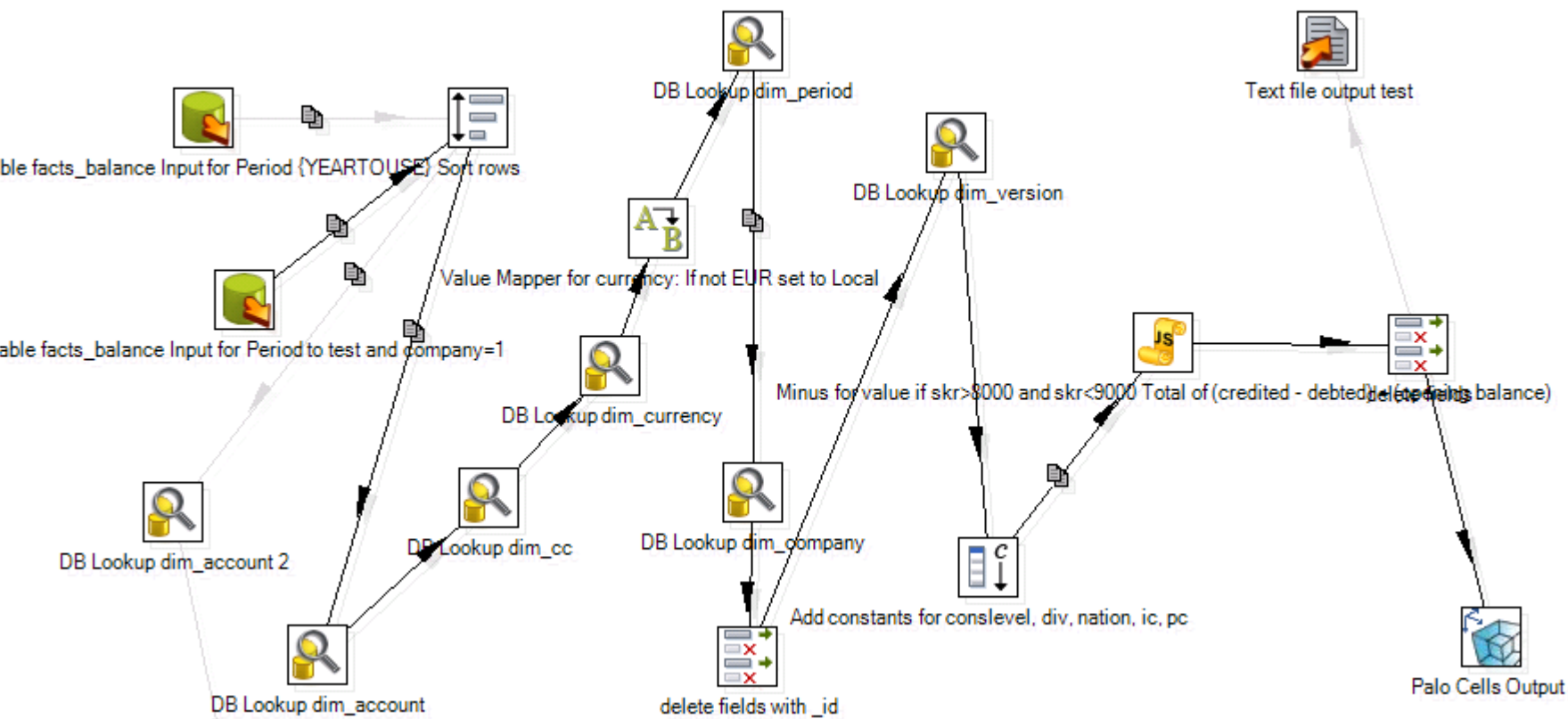
- public.dim_period
- public.dim_version
- public.facts_gl_pl_accounting
- public.facts_gl_pl_balance

The main window shows the table 'dw-h-oebif.public.facts_gl_pl_accounting' with the following data:

cy_id	conslevel_id	company_ic_id	division_id	nation_id	costcenter_id	costunit_id	profitcenter_id	date	amount_debted	amount_credited	account_contr
1	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	206560,1	
2	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	206560,1	
3	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	153177,77	
4	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	153177,77	
5	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	4000000	
6	3	1	1	1	1	1	1	01.01.2009 00:00:00	0	4143,15	



Starschema in OLAP Struktur wandeln, Lookup der Elemente, Laden nach Palo



PDI: Job Datev KaReWe: Nach Durchlauf - im Log nachvollziehbar - stehen die Daten in Palo für den Zugriff mit Excel bereit



Beispielhafte Log Ausgabe in PDI - Nachvollziehbarkeit und Unterstützung bei der Fehleranalyse:

```
Execution results
History Logging Job metrics
2011/02/24 07:28:41 - Delete fields - Felder löschen.0 - Finished processing (I=0, O=0, R=22296, W=22296, U=0, E=0)
2011/02/24 07:28:41 - Datenbanksuche Costcenter dim_cc Kostenstelle.0 - Finished processing (I=125, O=0, R=22296, W=22296, U=0, E=0)
2011/02/24 07:28:41 - Sequenz Nr. movement hinzufügen - Start bei 1.0 - Finished processing (I=0, O=0, R=22296, W=22296, U=0, E=0)
2011/02/24 07:28:41 - Data Validator for flag debted/credited 2.0 - Finished processing (I=0, O=0, R=22296, W=11559, U=0, E=10737)
2011/02/24 07:28:41 - Felder Umsatz - amount_credited.0 - Finished processing (I=0, O=0, R=11559, W=11559, U=0, E=0)
2011/02/24 07:28:41 - Felder Umsatz - amount_debted.0 - Finished processing (I=0, O=0, R=10737, W=10737, U=0, E=0)
2011/02/24 07:28:42 - Table facts_gl_pl_accounting Output.0 - Finished processing (I=0, O=22296, R=22296, W=22296, U=0, E=0)
2011/02/24 07:28:42 - Delete company and period in question, Import Data based on Datev txt files - Starting entry [SQL: Delete balances for ${company}, ${YEARTOUSE} PeriodMaxMin]
2011/02/24 07:28:42 - Delete company and period in question, Import Data based on Datev txt files - Finished job entry [SQL: Delete balances for ${company}, ${YEARTOUSE} PeriodMaxMin] (result=[true])
2011/02/24 07:28:42 - Delete company and period in question, Import Data based on Datev txt files - Finished job entry [Import-GL-fromDatevEBS-accounting-records] (result=[true])
2011/02/24 07:28:42 - Delete company and period in question, Import Data based on Datev txt files - Finished job entry [SQL: Delete accounting records for ${company}, ${YEARTOUSE} PeriodMaxMin] (re
2011/02/24 07:28:42 - Set Variables from Lookups on a per company base - Starting entry [Import Totals from MVZ and Calculate Totals from imported accounting records]
2011/02/24 07:28:43 - Import Totals from MVZ and Calculate Totals from imported accounting records - Starting job, loaded from repository : [/Martrade : Import Totals from MVZ and Calculate Totals from i
2011/02/24 07:28:43 - Import Totals from MVZ and Calculate Totals from imported accounting records - Starting entry [GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance]
2011/02/24 07:28:43 - GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance - Loading transformation from repository [GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance] in directory [/M:
2011/02/24 07:28:43 - GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance - Dispatching started for transformation [GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance]
2011/02/24 07:28:43 - GLtotals-from-facts_gl_pl_accounting-to-facts_gl_pl_balance - This transformation can be replayed with replay date: 2011/02/24 07:28:43
2011/02/24 07:28:43 - Table facts_gl_pl_balance.0 - Connected to database [PG82-Server mybiq-dwh] (commit=10000)
2011/02/24 07:28:43 - Generate Rows.0 - Finished processing (I=0, O=0, R=0, W=1, U=0, E=0)
2011/02/24 07:28:44 - Table facts_gl_pl_accounting all for period and company in question.0 - Finished reading query, closing connection.
2011/02/24 07:28:44 - Table facts_gl_pl_accounting all for period and company in question.0 - Finished processing (I=22296, O=0, R=0, W=22296, U=0, E=0)
```



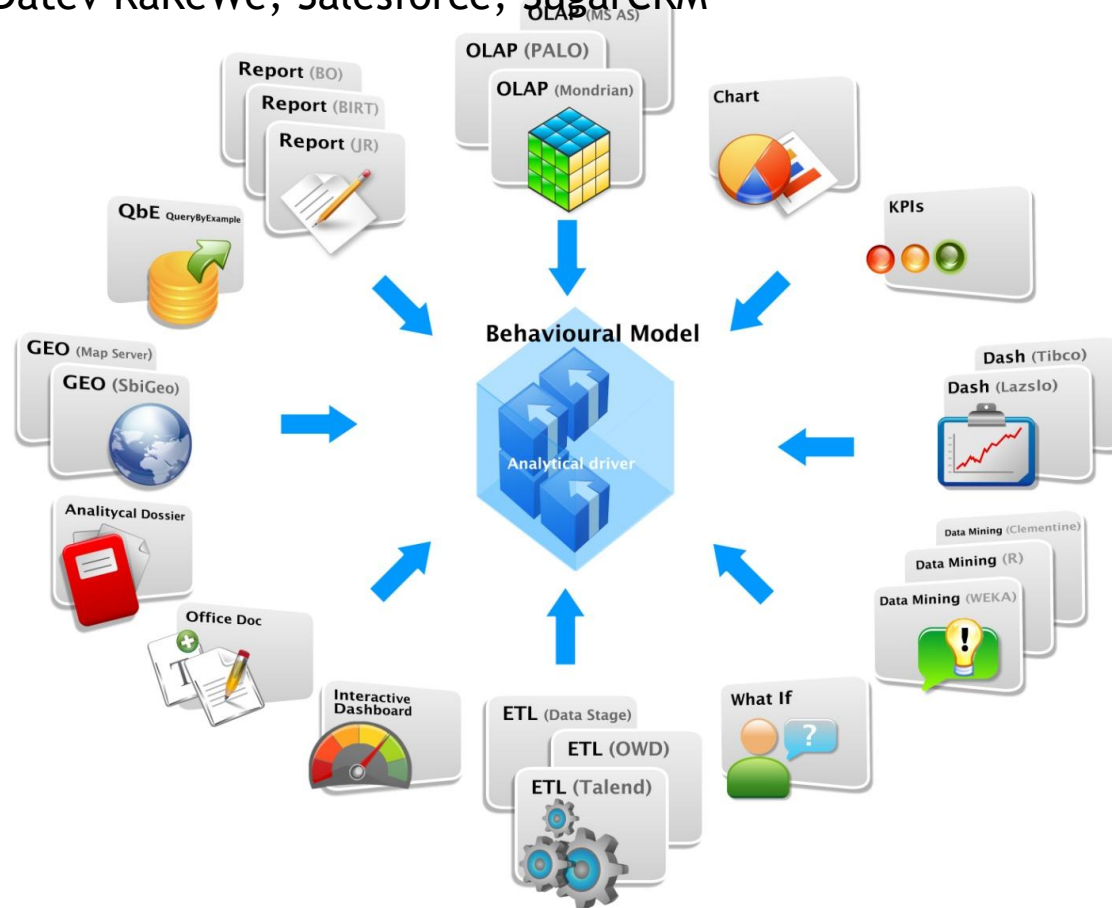
3A STRATEGY

Version 1.1, Feb. 2011 WS OSB
Hochschule KA

Webfrontend für den Endanwender mit Pivot, Charting, QBE, Dashboard



Datenquellen im DWH wie Palo, SAP BI, MS AS oder DBMS wie DB/2, Infobright, InfiniDB, MS SQL, MySQL, Oracle, PostgreSQL oder Applikationen wie SAP R/3, Navision, Datev KaReWe, Salesforce, SugarCRM





» Schätzung ist zu sportlich

- Change ist konstante Größe, wie so oft:

Die klassische “Beraterantwort” paßt also: Es gibt kein “richtig” oder es muß so sein ..., sondern es kommt darauf an ...

- Unterstützung durch IT schleppend, eigene Möglichkeiten und Know-How überschätzt, also: Puffer einplanen
- Datenqualität und Doku der OS läßt sich nicht einschätzen, Komplexität der Risiken nimmt zu, also: Transparenz so früh wie möglich, Kunden einbinden, Konsequenz:

» Sensibilisierung des Kunden hinsichtlich der Abhängigkeiten und der Komplexität

- Im richtigen Maße
- Zum jeweils richtigen Zeitpunkt
- Fakten sprechen lassen



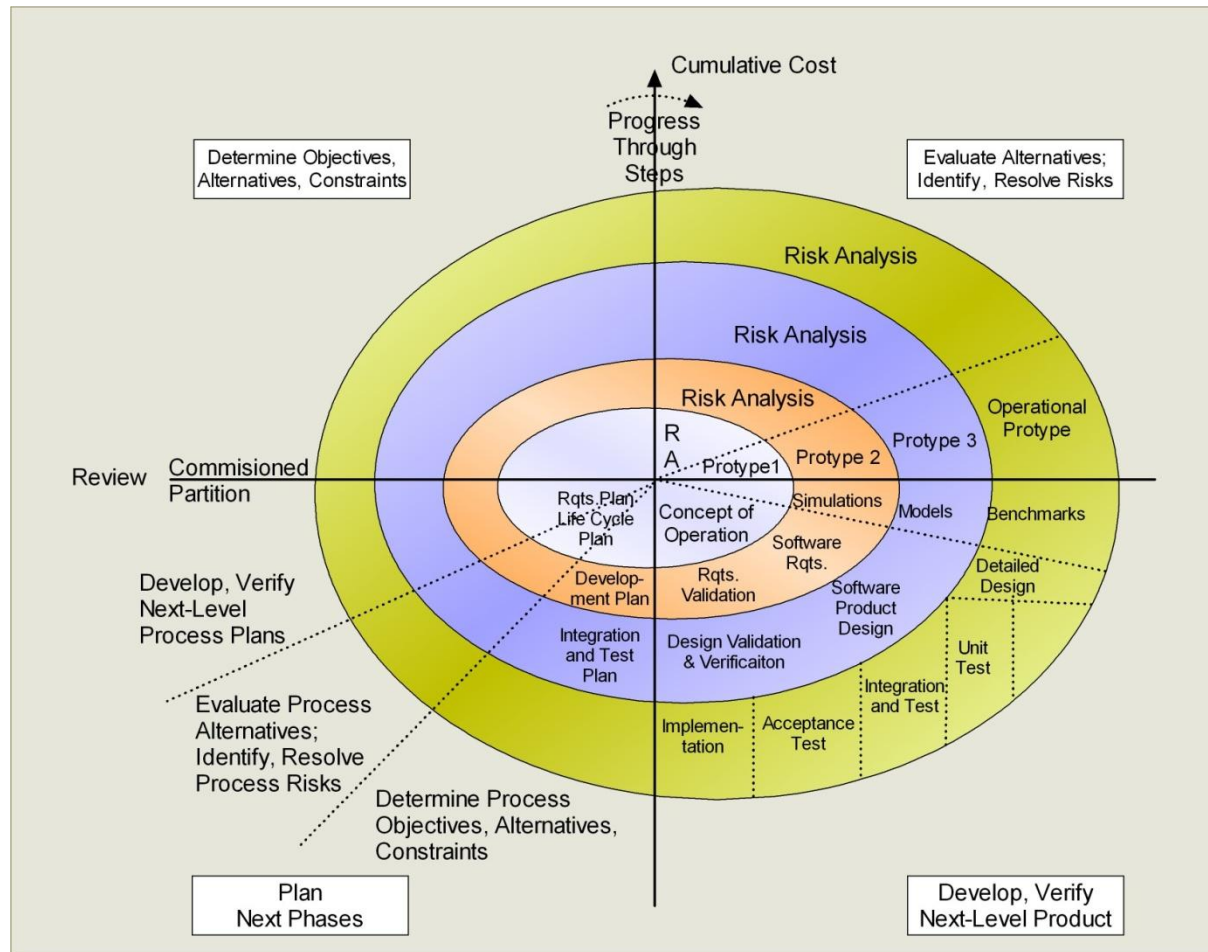
Projekt übergreifende Erkenntnisse - sprich: Lessons learnt (II)




- Einstellung auf den optimalen Level der Planungsgenauigkeit und Effektivität ist der Schlüssel zum Erfolg:
 - ➔ Permanente Anpassung bzw. Lerneffekt
 - ➔ Vgl. Lebenszyklus oder Regelkreis
- Konsequenz: “agile methology” ist eine notwendige Option, diese muß Bestandteil von Data Mart Kunden- und Projektmanagement sein!



Böhme - Spiralmodell





*¹Identify the impact of the
Schedule, Resources, S*

Projektbeispiele im Mittelstand: Einführung OSS für Business Intelligence

Komplexität transparent machen (ii)



<i>Assessment</i>	<i>Score</i>	<i>Rank</i>	<i>Value</i>
Human Resources	9	1	1.1
Communication	5	2	1.3
Procurement	10	3	3.8
Cost / Budget	10	4	5.0
Time / Schedule	4	5	2.5
Quality	8	6	6.0
Risk	9	7	7.9
Scope	7	8	7.0
Adjusted Complexity Indicator:			7,7



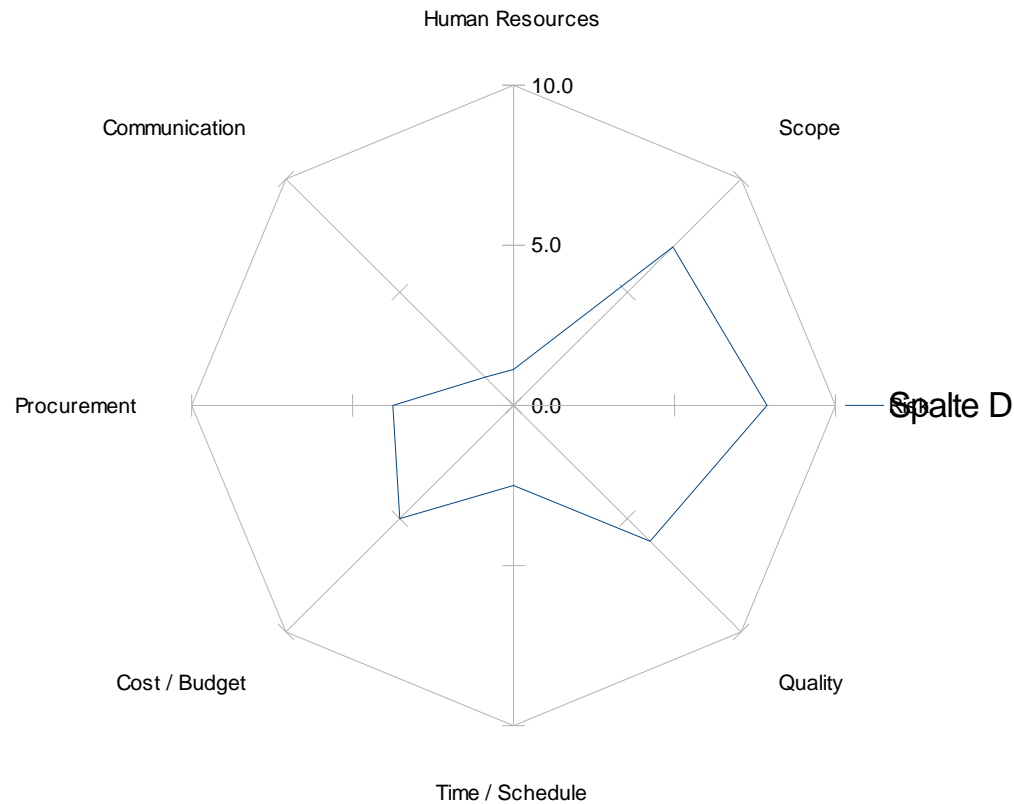
Komplexität transparent machen (i)

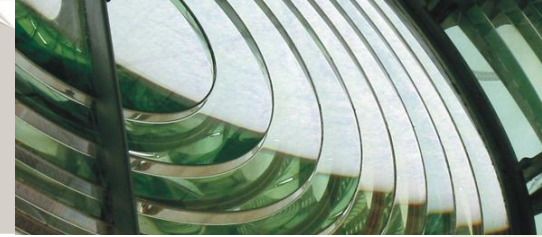


Definitions	Low Complexity Criteria	High Complexity Criteria
Scope	<ul style="list-style-type: none"> - Well understood scope elements - Each scope element is relatively independent 	<ul style="list-style-type: none"> - Scope elements are not clearly defined - Very complex or unclear relationships between scope elements.
Time / Schedule	<ul style="list-style-type: none"> - Minimal effort to be considered project by organization - All activities are independent 	<ul style="list-style-type: none"> - Maximal effort the organization would put in a single project - All activities are on critical paths
Cost / Budget	<ul style="list-style-type: none"> - Minimal budget to be considered stand-alone project - Uses a single cost account 	<ul style="list-style-type: none"> - Maximal budget that can be invested in a single project - Uses all potential cost accounts
Quality	<ul style="list-style-type: none"> - Well known business/activity area - QA processes defined and operational - QC procedures established and validated 	<ul style="list-style-type: none"> - New business activity - No existent QA processes - No existent QC procedures
Human Resources	<ul style="list-style-type: none"> - All in-house resources reporting to a single manager - Fully qualified resources (no training) 	<ul style="list-style-type: none"> - All external resources - No in-house technological experience for the project
Communication	<ul style="list-style-type: none"> - Single sponsor - Executed by a single organizational unit - Created for internal use only 	<ul style="list-style-type: none"> - Multiple sponsors - Multiple delivery participants - Multiple clients / impacted parties
Risk	<ul style="list-style-type: none"> - No visible risks 	<ul style="list-style-type: none"> - Highest risk level the organization would accept for a project, based on standard risk assessment
Procurement	<ul style="list-style-type: none"> - Only internal resources, no procurement - Formal procurement processes defined and validated 	<ul style="list-style-type: none"> - All external resources, each provided by a different external party - No formal procurement processes

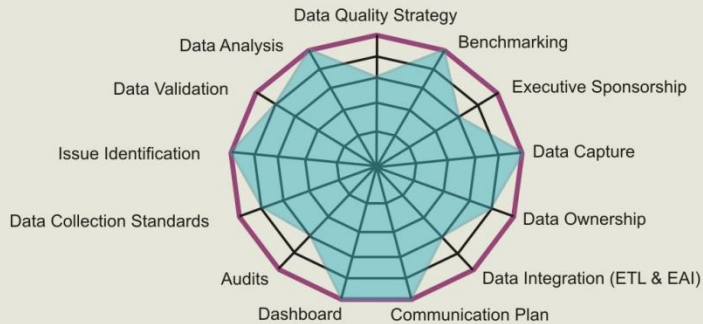


Komplexität transparent machen (iii)

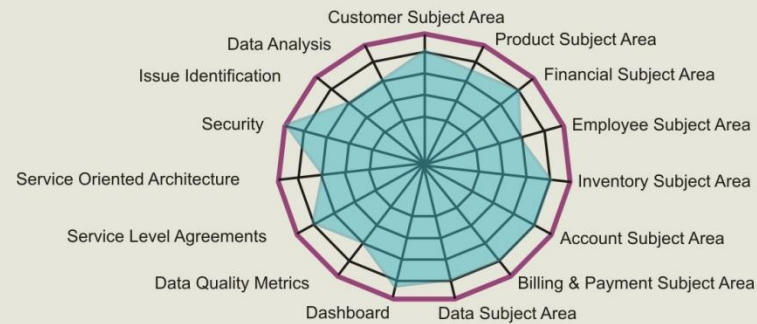




People/Organization



Compliance



Measurement



Policy



■ Capability ■ Agile



» Stader Saatzucht eG – Norddeutschland:



- Land- bzw. Agrarhandel im Raiffeisenverbund, 500 Mio. U
 - Saatgut, Dünger, Pflanzenschutz, Getreide- und Viehhandel
 - Baustoffe und Baumärkte
 - Tankstellen und Energiehandel
- 30 Standorte, 100.000 Artikel und Kunden
- 5 Mio. Verkaufstransaktionen (ohne Kassen) pro Jahr in Navision
 - Also überschlagsmäßig: 50 pro Artikel, 50 pro Kunde
 - 10 GB für Tabelle Artikelposten
 - 300 GB Datenvolumen
 - Nachtläufe und Backup dauern zu lange
 - Maximale Auslagerungsmöglichkeiten ausgereizt, derzeit 2 Jahre

» Konzept:

- Auslagerung - hochperformante spaltenbasierte, in-Memory Datenbank
- Frage: Ohne Aufwand bei Tools und Know-how - SQL nutzbar?





➤ Analyse der verschiedenen Angebote

- Open Source Lösungen
- Kommerzielle Lösungen
- Praktikabilität und guter Support im Forum -> Entscheidungsgrundlage

➤ Start: Auswertung Batch Lauf "Warenbestandslisten"

Code	WBL-Pos	LG	Bereich	Bereich1	Anf-BST	EK	Zu	VK	Ab	BST	WBL-BST	Verb	DB	VK_Me
B0	130		B		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
D-D0	190		D-D		0,00	-8.271,21	9.118,22	201,78	0,00	4,87	0,00	-168,98	-645,23	1.306,00
D-P0	870		D-P		0,00	-7.916,46	18.462,14	10.908,64	0,00	-5.241,67	0,00	-10.212,96	-4.879,68	191,83
D-S0	90		D-S		0,00	-320,00	-234,55	-320,00	0,00	0,00	0,00	0,00	234,55	0,00
E-LGTW0	120		E-LGTW		0,00	19.513,72	-21.229,63	1.630,86	0,00	227,42	0,00	1.105,33	3.346,77	804,00
E-TST0	150		E-TST		0,00	31.352,10	-31.612,04	321,12	0,00	0,00	0,00	0,00	581,06	0,00
F0	320		F		0,00	1.029,97	-802,16	810,50	0,00	0,00	0,00	-705,35	582,69	8,00
Gem0	70		Gem		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
H0	4060		H		7.998,18	16.826,21	-17.604,18	610,78	0,00	7.963,21	7.963,21	-470,76	1.353,78	60,30
K0	290		K		0,00	5.388,75	0,00	6.337,39	0,00	0,00	0,00	-6.199,18	948,64	66.253,00
S0	120		S		0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00	0,00
B1	241		B		1.950,75	11.742,14	-3.221,16	6.287,18	0,00	1.077,13	1.077,13	-5.047,98	-3.107,42	1.597,00
D-D1	1201		D-D		163.749,95	1.128.872,04	-147.968,09	1.173.805,75	0,00	164.042,61	141.645,68	-1.069.801,17	170.797,53	10.252.008,89
D-P1	5151		D-P		118.206,56	511.341,91	37.266,84	600.611,69	0,00	115.272,30	117.644,23	-548.622,20	51.440,61	13.309,86
D-S1	1241		D-S		30.821,59	37.577,35	-3.958,39	45.114,87	0,00	30.371,47	30.209,01	-36.836,78	10.883,33	29.112,34
E-LGTW1	451		E-LGTW		8.905,34	21.571,84	-6.257,38	10.767,83	0,00	13.622,08	13.622,08	-8.196,76	170,11	8.294,34
E-TST1	21		E-TST		2.293,25	-125.750,29	777.540,10	669.236,37	0,00	19.452,27	19.500,52	-623.567,13	34.653,83	697.128,03
F1	12021		F		1.490.379,84	12.852.384,64	1.666.794,90	17.643.539,51	0,00	1.077.032,70	792.169,76	-16.677.259,80	2.426.149,90	116.224.487,00

➤ Bis dato aus Navision ermittelt, nach Excel kopiert – nur einmal jährlich Auswertung der 11 Mio. Datensätze möglich





» Lösung:

- Separates Data Warehouse (DWH)
 - 2 PC's aus Mediamarkt oder Saturn ... Stückpreis 450 Euro
 - Linux / Windows und Installation der OSS OE**bif**
- Spiegeln der Navision Daten auf 32-bit PC mit Windows
- Extraktion Navision Daten mit Pentaho und NaviX
- Datenbereinigung ("Cleansing") und Mapping
- Laden der Daten nach Infobright 64-bit PC

» Nutzen:

- Auswertung und Aggregationen wöchentlich - separate PC's
- OLAP/Excel und Webfrontend für benutzerfreundliche Auswertungen
- Möglichkeit der Erweiterung als Dashboard

» Aufwand in Manntagen für Artikelposten mit Webfrontend:

- 8 Manntage
- Software Lizenzkosten: Keine
- Laufende Kosten: Nur bei Änderungen auf der ERP-Seite ggf. entsprechend Anpassung notwendig



PDI: Auslesen der Navision bzw. MS Dynamics NAV Tabelleninhalte per NaviX



in schemas
server
cluster sche

Navigation SSZ Artikelposten

Text file output exp-navision-artikelposten \${DATADIR} \${SEP} \${ENCL}

NaviX Table Output

Step name: Navigation SSZ Artikelposten

Navigation Connection | Navigation Table | Navigation Table Fields

No	Name	Alt. Name	Type	L...	S..	Class	Active	Condition
1	Lfd. Nr.	Lfd. Nr.	Integer	4	4	Normal	true	
2	Artikelnr.	Artikelnr.	Code	20	22	Normal	true	
3	Buchungsdatum	Buchungsdatum	Date	4	4	Normal	true	
4	Postenart	Postenart	Option	4	4	Normal	true	
5	Herkunftsnr.	Herkunftsnr.	Code	20	22	Normal	true	
6	Belegnr.	Belegnr.	Code	20	22	Normal	true	
7	Beschreibung	Beschreibung	String	50	51	Normal	false	
8	Lagerortcode	Lagerortcode	Code	10	12	Normal	true	
9	Lagerbuchungsgruppe	Lagerbuchungsgruppe	Code	10	12	Normal	false	
10	Herkunftsbuchungsgruppe	Herkunftsbuchungsgruppe	Code	10	12	Normal	false	
11	Menge	Menge	Decimal	12	12	Normal	false	
12	Restmenge	Restmenge	Decimal	12	12	Normal	false	
13	Fakturierte Menge	Fakturierte Menge	Decimal	12	12	Normal	true	
14	Stückpreis	Stückpreis	Decimal	12	12	Normal	false	
15	Einstandspreis	Einstandspreis	Decimal	12	12	Normal	false	
16	Betrag	Betrag	Decimal	12	12	Normal	true	
17	Rabattbetrag	Rabattbetrag	Decimal	12	12	Normal	false	
18	Verk./-Einkäufercode	Verk./-Einkäufercode	Code	10	12	Normal	true	
19	Benutzer ID	Benutzer ID	Code	20	22	Normal	false	
20	Herkunftscode	Herkunftscode	Code	10	12	Normal	true	

Refresh
Deactivate
Activate
Select All
Unselect

Provided by 3A-Strategy | OEbif

Check Entries Apply OK Cancel



DWH - Extraktion Navision bei Stader Saatzucht eG

Spoon - [localhost pdi version 4] Control NAV export - create and load DWH SSZ

File Edit View Action Tools Help

Perspective: Data Integration

Control NAV export - create and load DWH SSZ DWH Aggregate Artikelposten

100%

START

Set variables for db connection dwhdatabase and delimiters

Set variable ARTPOSTEN

Export Data from Navision

Insert records from NAV export files to Infobright

DUMMY

Create DWH based on Navision structures 2

Import DWH to DWH by files Exported from Navision 2

Import DWH to DWH by files Exported from Navision

Create DWH based on Navision structures

Shell for starting infobright

Set variables for db connection dwhdatabase on Infobright

Execution results

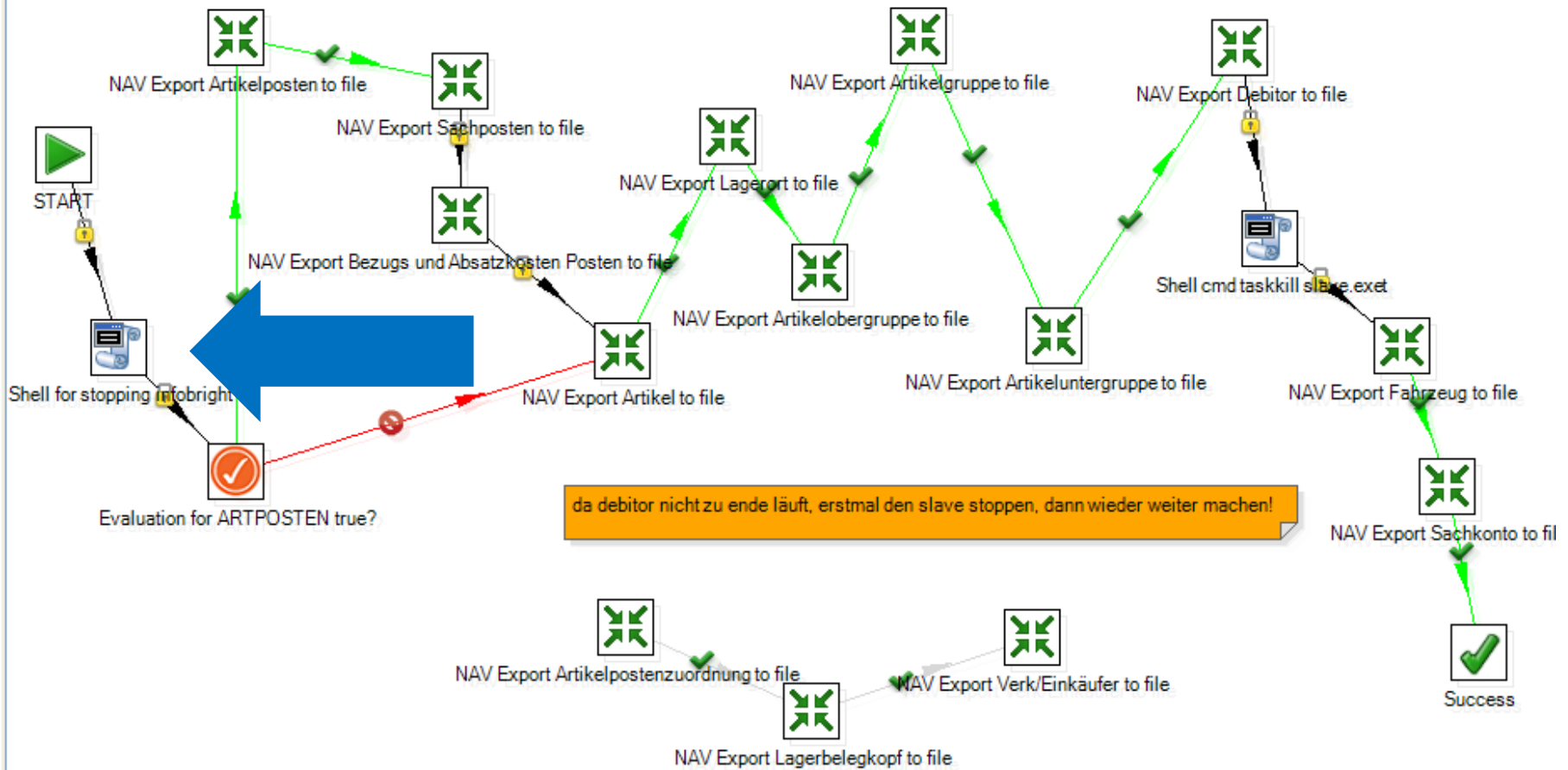
History Logging Job metrics

2011/01/25 10:07:55 - Text file output exp-navision-artikelposten.0 - linenr 11500000
2011/01/25 10:08:04 - Navision SSZ Artikelposten.0 - Finished processing (I=11515545, O=0, R=0, W=11515545, U=0, E=0)
2011/01/25 10:08:04 - Text file output exp-navision-artikelposten.0 - Finished processing (I=0, O=11515545, R=11515545, W=11515545, U=0, E=0)
2011/01/25 10:08:04 - Export Data from Navision - Starting entry [NAV Export Sachposten to file]
2011/01/25 10:08:04 - NAV Export Sachposten to file - Loading transformation from repository [NAV Export Sachposten to file] in directory [/SSZ]
2011/01/25 10:08:05 - NAV Export Sachposten to file - Dispatching started for transformation [NAV Export Sachposten to file]
2011/01/25 10:08:05 - Navision SSZ Sachposten.0 - Starting...
2011/01/25 10:08:05 - NAV Export Sachposten to file - This transformation can be replayed with replay date: 2011/01/25 10:08:05
2011/01/25 10:10:12 - Text file output exp-navision-sachposten.0 - linenr 500000
2011/01/25 10:12:07 - Text file output exp-navision-sachposten.0 - linenr 1000000
2011/01/25 10:14:00 - Text file output exp-navision-sachposten.0 - linenr 1500000
2011/01/25 10:15:50 - Text file output exp-navision-sachposten.0 - linenr 2000000
2011/01/25 10:17:45 - Text file output exp-navision-sachposten.0 - linenr 2500000
2011/01/25 10:19:32 - Text file output exp-navision-sachposten.0 - linenr 3000000
2011/01/25 10:21:29 - Text file output exp-navision-sachposten.0 - linenr 3500000
2011/01/25 10:23:24 - Text file output exp-navision-sachposten.0 - linenr 4000000
2011/01/25 10:25:19 - Text file output exp-navision-sachposten.0 - linenr 4500000
2011/01/25 10:27:15 - Text file output exp-navision-sachposten.0 - linenr 5000000
2011/01/25 10:29:12 - Text file output exp-navision-sachposten.0 - linenr 5500000
2011/01/25 10:30:00 - Text file output exp-navision-sachposten.0 - linenr 6000000

Projektspeispiele im Mittelstand: Einführung OSS für Business Intelligence

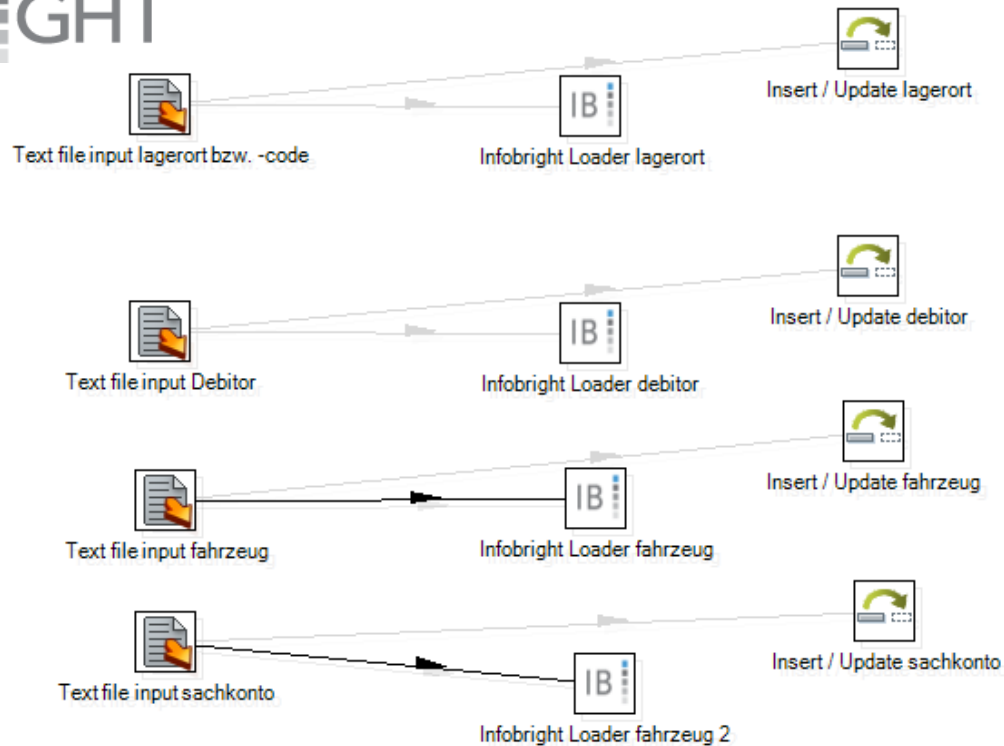


DWH - Extraktion Navision bei Stader Saatzucht eG



DWH - Laden der Daten in Infobright

INFOBRIGHT





- Auswertung mit SQL möglich
- Keine normale RDBMS hat die Performance einer spaltenbasierten, analytischen Datenbank wie Infobright
- Bsp. Count auf die Tabelle Artikelposten – Zählen pro Artikel, Beschränkung auf 7700 Einzelsätze/Artikel von 11 Mio.:
 - 15,81 Minuten in MySQL, also klassisches RDBMS
[Executed: 08.01.11 16:19:02 MEZ] [Execution: 948531/ms]
 - 0,31 Minuten oder 18,84 Sekunden
[Executed: 08.01.11 16:18:12 MEZ] [Execution: 18844/ms]

INFOBRIGHT

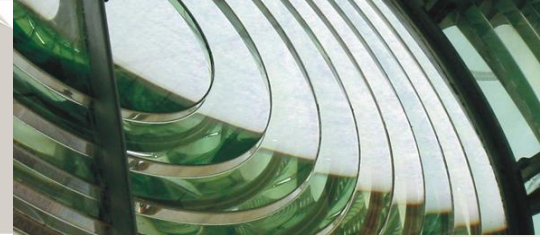


DWH -Daten Auswerten in Infobright hier mit COUNT auf Navision Tabelle Artikelposten

```
SELECT
    `Nr.`,
    Artikel.Beschreibung,
    count(*)
FROM dwh_nav.artikelposten join dwh_nav.artikel
WHERE `Artikelnr.`=`Nr.` AND `Nr.` < 54050000000
GROUP BY `Artikelnr.` ORDER BY `Artikelnr.`
```

Nr.	Beschreibung	count (*)
1010	Schwein Handelsklasse U	5
1020	Schwein Handelsklasse R	8
1050	Schwein	8312
1051	EG Schweine Marke	4423
1052	EG Schweine Programm	682
1060	Schwein Notschl./Tot	597
...		
54040000194	Bauder Pur 025 G 60 mm	31
54040000196	Bauder/Selthaan diverse	169
...		
54040000351	Extrudierter Hartschaum	483
54040000352	Extrudierter Hartschaum	291
54040000353	Extrudierter Hartschaum	401

INFOBRIGHT



» Strukturen & Daten für DWH in MySQL bzw. Infobright:

- Gleiches SQL Interface, geringe Lernkurve, OSS Variante für Beginn
- Super performante Aggregierung mit Low Cost PC mit Historisierung auf Belegebene
- Palo bietet kein echtes Drill Through oder Datamining, Datenvolumina limitiert

» Nutzung von Palo mit Excel:

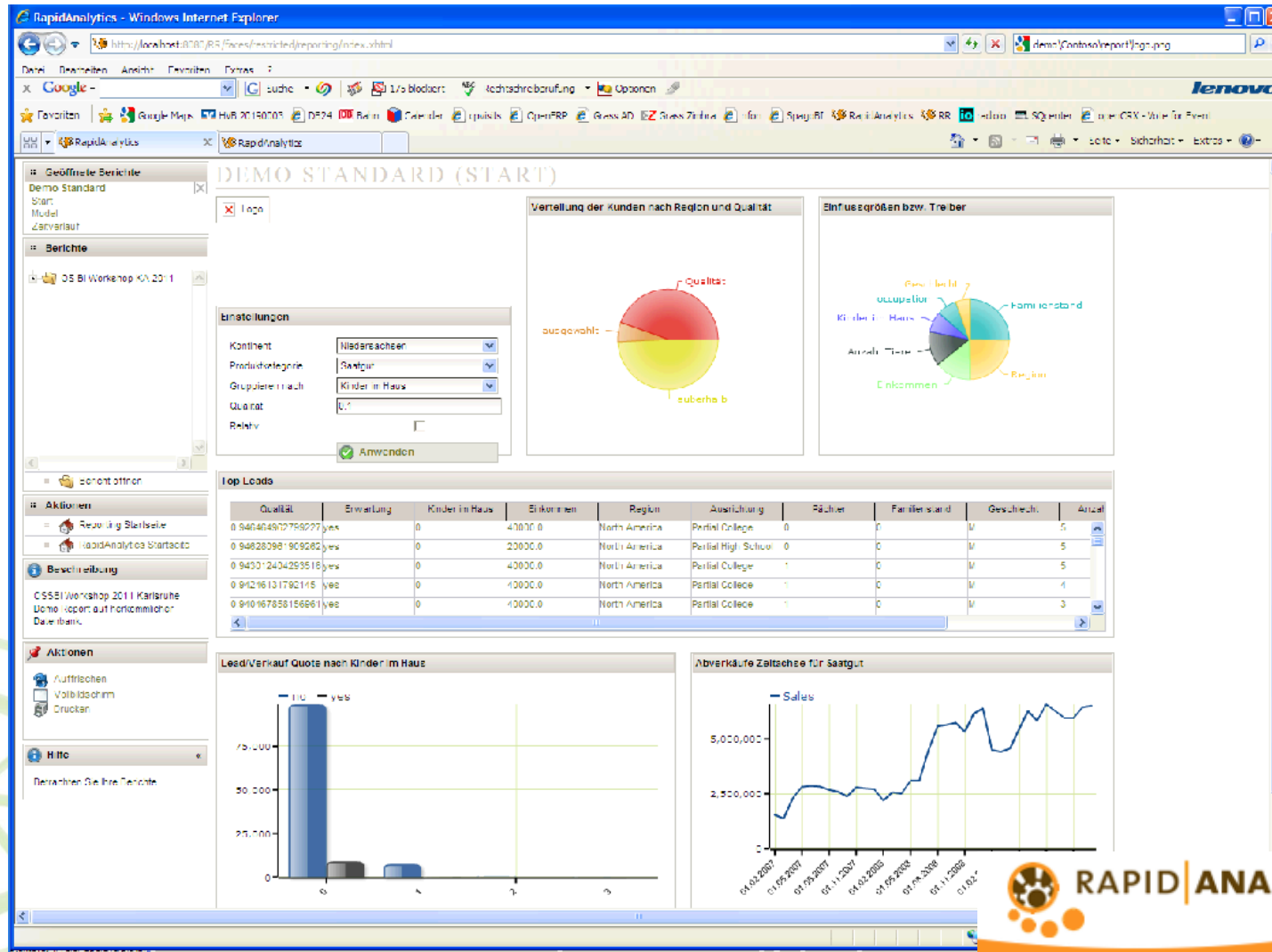
- Palo als abstrakte Controlling Ebene, etwa Planung strategisch und auf weiteren operativen Ebenen
- Vertriebsdaten angereichert um Finanzdaten zur Planung auf Kundenebene

» Webportal mit RapidAnalytics für Infobright:

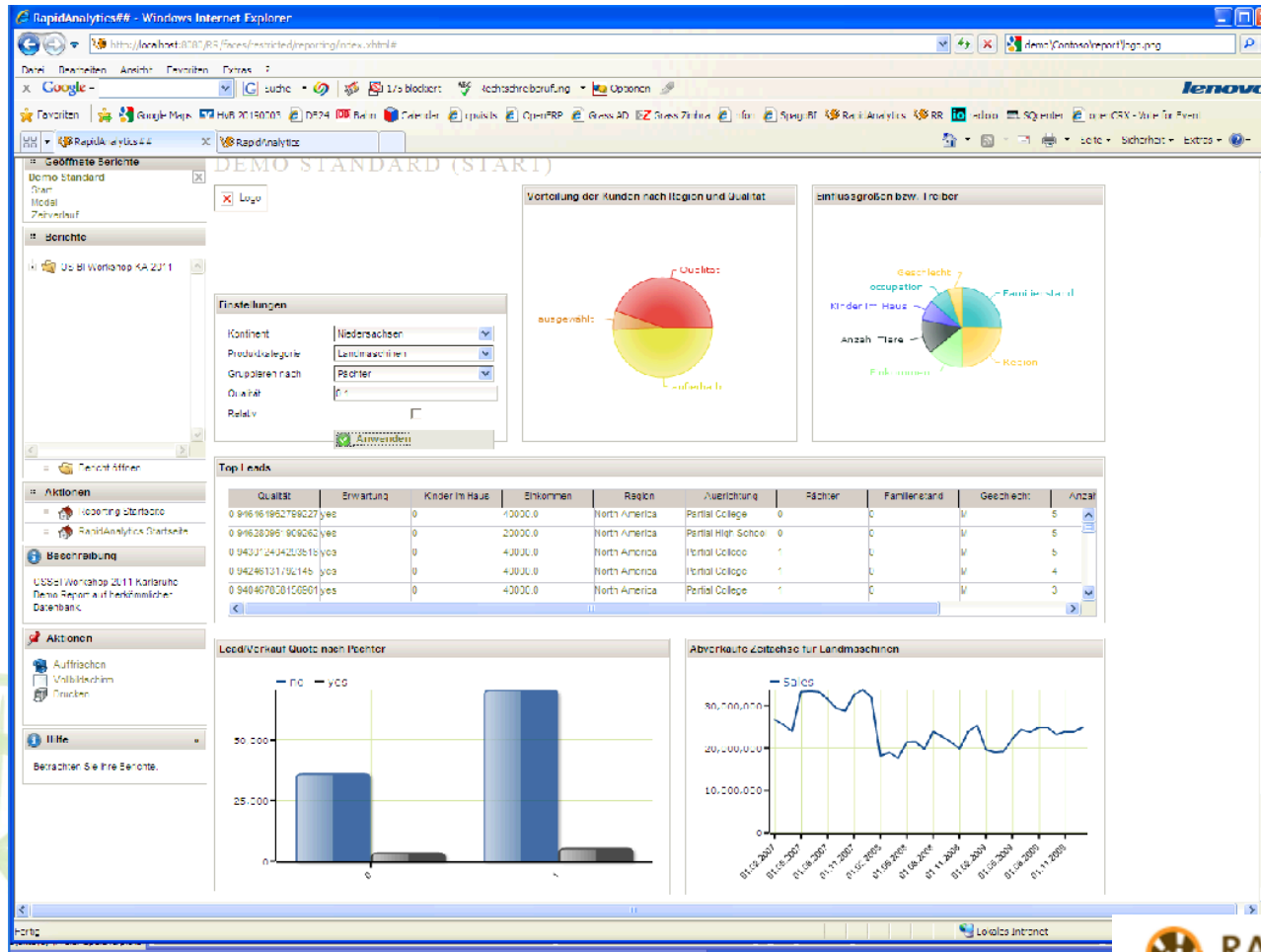
- Einheitliche GUI für alle Anwender
- Breiter Funktionsumfang inkl. Datamining für Warenkorbanalyse



Webportal mit Datamining Funktionen: RapidAnalytics - Einsteigerversion als OSS



Webportal mit Datamining Funktionen: RapidAnalytics - Einstiegsversion als OSS



Vielen Dank für Ihre Aufmerksamkeit!



- » Waren Anregungen für Sie dabei?
- » Haben Sie Fragen zu meinen Argumenten?
- » Haben Sie ähnliche Erfahrungen gemacht?
- » Meine Kontaktdaten haben Sie ja ...





Kontakt

Für Fragen und weitere Informationen wenden Sie sich bitte an:

Alexander Pohl

CEO / Geschäftsführer

aPohl@3A-Strategy.com

Büro: +49 (89) 2190.985-58

Mobile: +49 (160) 584.2020

Fax: +49 (40) 85.37.82-705

vertreten durch die Gründer:

3A-Strategy Global Advisory GmbH

Hamburg - München

Lilienstraße 11

20095 Hamburg

Infanteriestraße 11e Suite 4

80979 München