

Best Practices: BI mit Open-Source-Tools

Alf Hellmund - GIUA 2009

Agenda

Einleitung

- Vorstellung & Motivation
- Vorteile

Best Practices

- Architektur & Entwurf
- Datenmodellierung
- ETL
- Reporting
- Service

Fazit

- Zusammenfassung
- Fragen?

Wer sind wir? Was machen wir?



- Namensklärung:
decerno (Entscheidung) und **opus** (Fabrik)
- Beratung / Consulting mit Schwerpunkt im Bereich Datenbanken & DWH
- Produktentwicklung
- Erfahrung in über 30 Projekten in den Bereichen Analyse, Architektur, Datenversorgung, Entwicklung, Reporting und Re-Design

Vorteile von BI mit Open-Source



- Prozesse zur systematischen Analyse (Sammlung, Auswertung und Darstellung) von Daten
- Einfache Integration heterogener Datenquellen
- Verbesserung der Datenqualität
- Historische Betrachtung von Entwicklungen
- Verdichtung der Daten
- Ausgefeilte Analyse und Reporting möglich
- Keine Lizenzkosten, Planungssicherheit
- Eigene Weiterentwicklungen
- Geringer Schulungsaufwand

Agenda

Einleitung

- Vorstellung & Motivation
- Vorteile

Best Practices

- Architektur & Entwurf
- Datenmodellierung
- ETL
- Reporting
- Service

Fazit

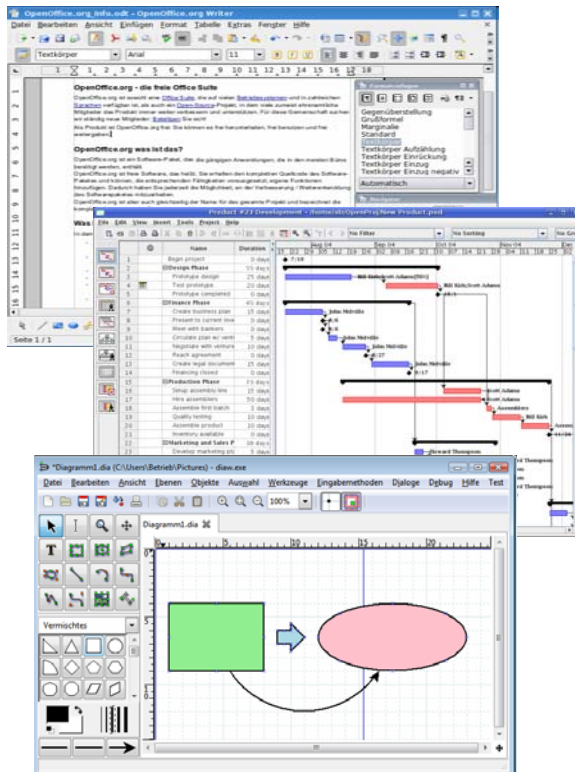
- Zusammenfassung
- Fragen?

Architektur & Entwurf: Probleme



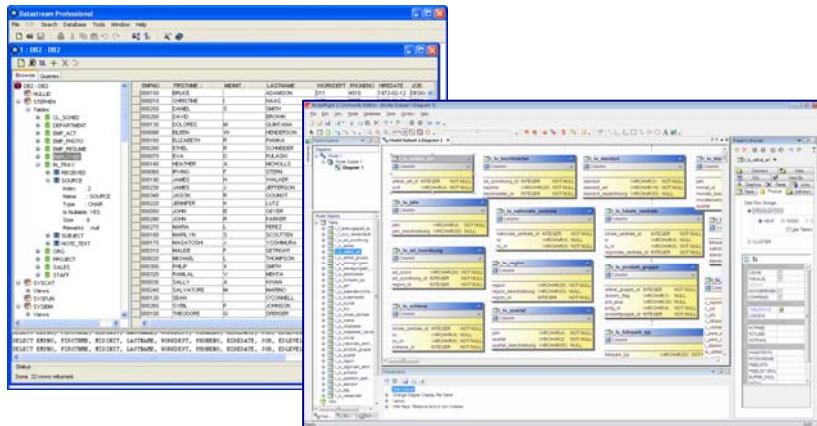
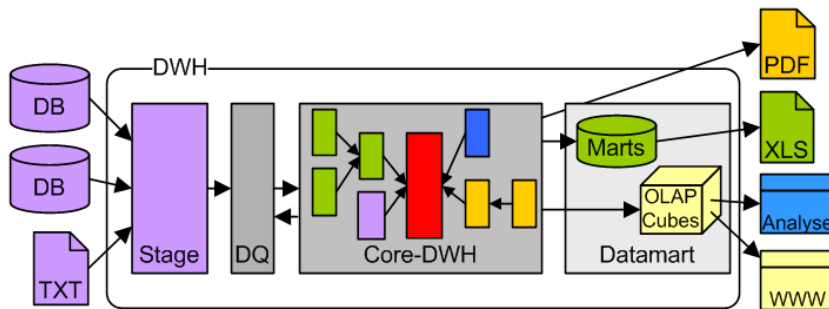
- Keine genauen oder übergroße Zielvorstellungen
- Fehlende Kommunikation zwischen Fachbereich und IT
- Mangelndes Verständnis für Abhängigkeiten
- Fehlende Dokumentation der Geschäftsprozesse
- Widersprüchliches Regelwerk
- Fehlende Motivation
- Fehlende Unterstützung durch die Geschäftsleitung

Architektur & Entwurf: Lösungen



- Einbindung der Projektbetroffenen und Entscheider
- Gemeinsame Vorgehensweisen für Prozesse (Analyse, Entwicklung, Test, Betrieb)
 - *trans|pa|rent* (durchsichtig; durchschaubar)
 - *effi|zi|ent* (wirksam; wirtschaftlich)
- Evolutionäre Entwicklungszyklen möglich, da keine hohen Anfangsinvestitionen
- Scope offen halten
- Aufsplittung des Projektziels in Meilensteine
- Erstellung und Verwendung von Testkonzepten
- Lessons-Learned
- Open Office, OpenProj und DIA

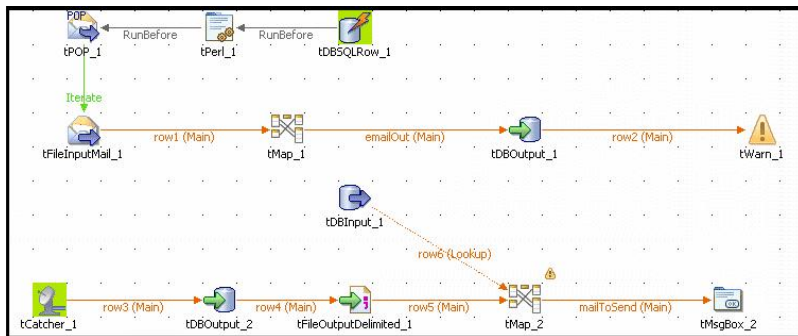
Datenmodellierung



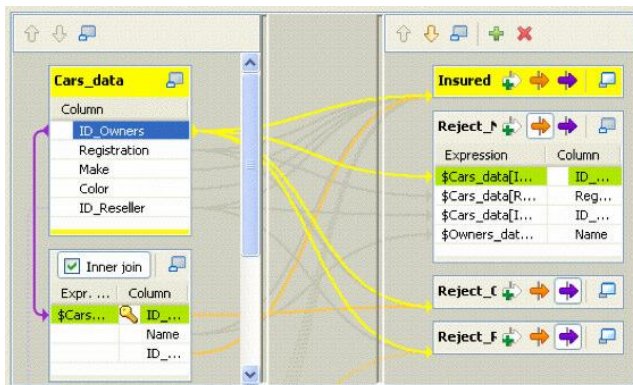
- Planung eines DWH in Schichten:
 - Daten-Stage (ggf. mit DQ)
 - Core-DWH als Star- bzw. Snowflake-Schema
 - OLAP-Cubes & Datamarts
- Nutzung von Case- & Abfrage-Tools:
 - Reverse Engineering
 - Leichte Wartbarkeit
 - Automatische Codegenerierung, Syntaxprüfung
 - Zusammenhänge der Tabellenstruktur sind übersichtlicher und dokumentiert
 - DB Browser, Datastream Pro

Best Practices: BI

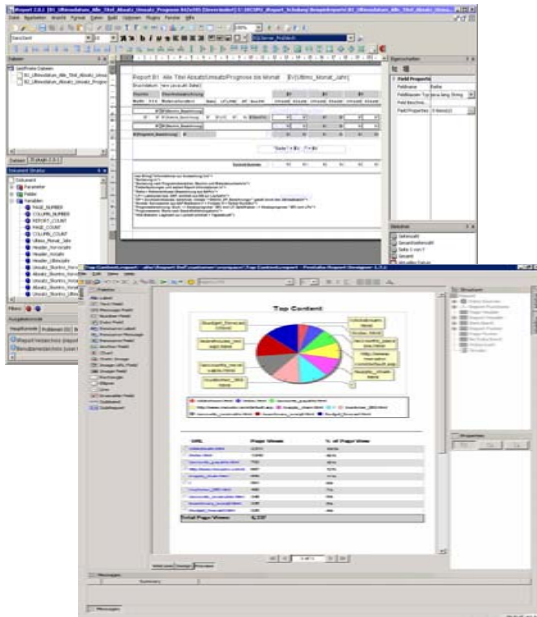
ETL



- Schnelle und effiziente Erzeugung der Prozesse durch JasperETL (Talend) oder Pentaho (Kettle)
- Sehr gute Modifizier- und Erweiterbarkeit durch Nutzung erprobter Transformationen
- Leichter Zugriff auf heterogene Quellen
- Umfangreiche Verbesserung der DQ
- Konzentration auf fachliche Anforderungen anstatt auf Bugfix von Codezeilen z.B. bei Historisierung
- Hohe Skalierbarkeit, Monitoring
- Vereinigung der Daten in einem „Single Point of Truth“
- Versionierung der Änderungen
- Aufbau komplexer Workflows

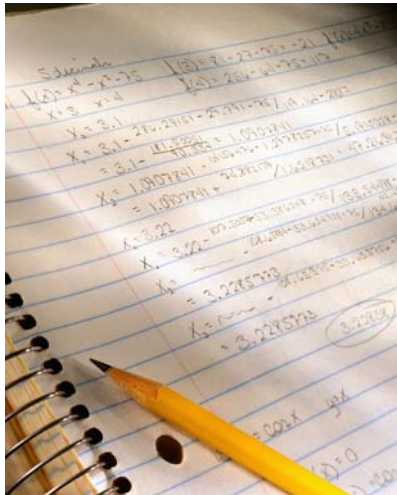


Reporting



- Einfacher und schneller Entwurf von Reports mittels Generatoren (iReport, Pentaho Reporting)
- Direkter Excel-Zugriff mittels Datamarts
- Direkte Datenauswertung und –analyse über Cubes
- Ad-Hoc-Reporting: Konzentration auf die Fachlichkeit
- Webbasierte Zugriff auf Reports und Dashboards
- Einfache Austauschbarkeit durch standardisiertes Datenmodell

Service



- Ein Projekt endet nicht mit der Produktivsetzung!
- Durch den Einsatz der BI-Werkzeuge Synergieeffekte im Betrieb und für die Weiterentwicklung nutzen
- Leichte Erweiterbarkeit und minimaler Testaufwand durch schnelles Eingrenzen der relevanten Bereiche
- Automatische Code-Generierung bei Änderungen vermeidet Fehler
- Vereinfachte Dokumentation auch nach einer Aktualisierung
- Planungssicherheit durch Open-Source
- Reduzierter Schulungsaufwand durch transparente Prozesse

Agenda

Einleitung

- Vorstellung & Motivation
- Vorteile

Best Practices

- Architektur & Entwurf
- Datenmodellierung
- ETL
- Reporting
- Service

Fazit

- Zusammenfassung
- Fragen?

Zusammenfassung



- Klärung der fachl. & techn. Aufgabenstellung vorab
- Case-Tools unterstützen bei Planung, Umsetzung, Dokumentation & Wartung
- DB-Modellierung-Tools
- Sehr ausgereifte Open-Source-Tools für ETL und Reporting verfügbar
- Eigene Modifikationen dank Quellcode möglich
- Unterstützung durch Hersteller und Community
- Keine Lizenz-, nur Entwicklungs- und Wartungskosten

Best Practices: BI

Haben sie Fragen?

Vielen Dank für Ihre Aufmerksamkeit!