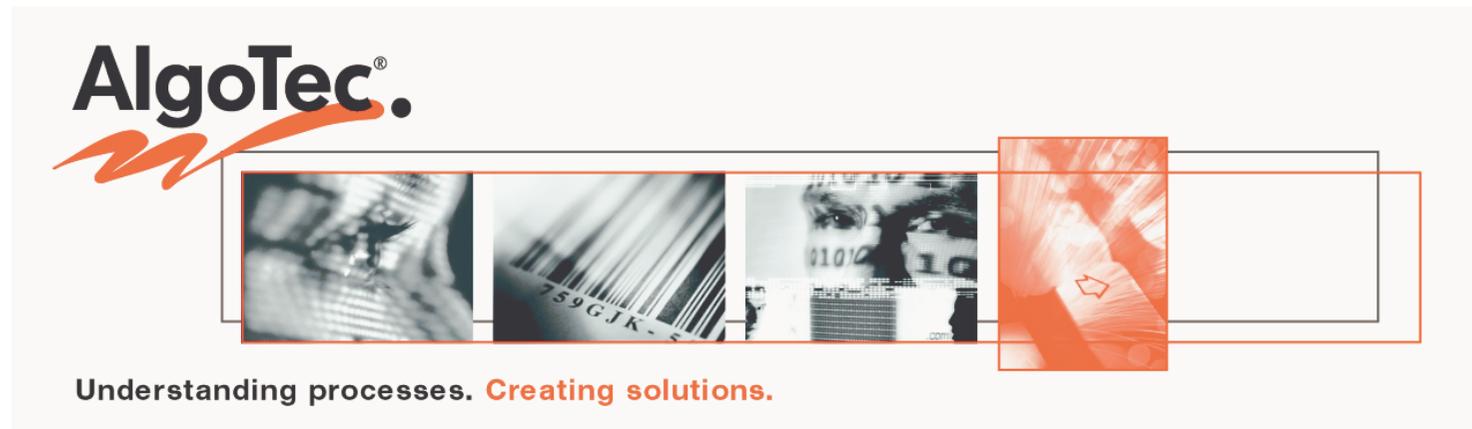


# Datenmigrationen, oder die Bombe mit der langen Zündschnur

Martin Weikinger, AlgoTec GmbH





# Vortragsstruktur



- Themenabgrenzung und Motivation
- besondere Migrationsprobleme und –  
gefahren
- 3 goldenen Migrationsregeln
- 4-Phasen-Migrationsprozess
- Zusammenfassung / Fazit



## Abgrenzung und typisches Szenario



- Datenmigrationen, jedoch keine Anwendungs- oder Systemmigrationen
- Anwendungs- oder OS-Spezifika unberücksichtigt
- typisches Szenario
  - Einführung eines neuen Systems
  - Ablösung eines oder mehrerer Altsysteme
  - „Impfen“ des neuen Systems mit Bestandsdaten



- Gute Datenmigrationen sind wichtig weil sie
  - die versprochenen Vorteile des neuen Systems erst wirklich nutzbar machen
  - mit maximaler Datenqualität starten: die besten, präzisesten und saubersten Daten importieren zur Verbesserung der Business Intelligence
  - legale und auditive Anforderungen einhalten
  - reibungslosen Projektablauf fördern





## Was Datenmigrationen schwierig macht ... [1/2]



- „private“ lokale IT-Lösungen des Fachbereichs
  - die IT-Affinität beim Fachbereich ist gewachsen
  - prozesskritische Datenbestände (oft kleinere DBs oder Spreadsheets) entstehen dezentral ohne dass die zentrale IT davon etwas erfährt
  - Qualität und Schlüssel uneinheitlich (nix UWDM / connectIT)
  - potentielle Dokumentationsdefizite und der Verlust des „geistigen Vaters“ drohen
  - Daten oft spät erst als migrationsrelevant erkannt



## Was Datenmigrationen schwierig macht ... [2/2]



- die Bedeutung von Datenfeldern kann über die Zeit und Unternehmensteile hinweg variieren
- Experten für Neusystem kein Problem, aber für Altsystem ?
- verstreutes Wissen: in der heutigen komplexen Unternehmenswelt gibt es keine einzelne Person, die weiß wie das Unternehmen arbeitet





## ... und warum sie geradewegs gefährlich sein können [1/2]



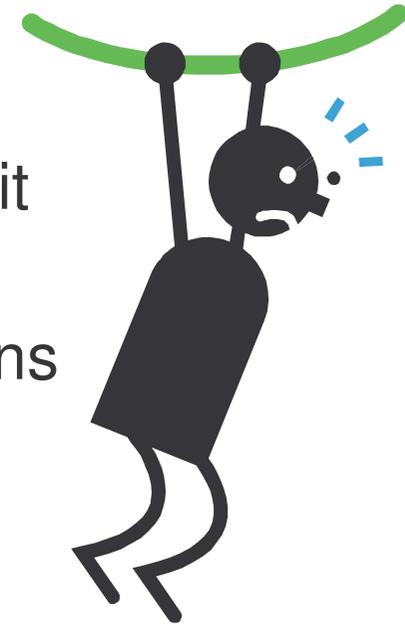
- Aschenputteleffekt: die Datenmigration ist nur ein (scheinbar hässliches kleines) Teilprojekt innerhalb eines größeren Gesamtprojekts
  - einseitiger Fokus auf das neue System
  - die besten Mitarbeiter setzen das neue System auf: „Unser zweiter Sturm wird die Migration schon schaffen ...“
  - Putzfrauensyndrom: Wenn Du einen perfekten Job machst, wird's keiner merken, mach's falsch und jeder kennt Dich !
  - i.d.R. wenig karrierewirksam



## ... und warum sie geradewegs gefährlich sein können [2/2]



- die späte Positionierung der Migration innerhalb der Projektphasen
  - „Wir haben jetzt doch genug um die Ohren, dafür ist später noch genug Zeit ...“
  - ... wo oft Budget- und Zeitplanungen ins Rutschen geraten
    - Folgen für die Datenqualität
  - Bombe mit langer Züandschnur: im Projekt ein „Single-Point-Of-Failure“ mit später Auswirkung





# 1. goldene Migrationsregel



- eine Datenmigration ist ein fachliches und kein technisches Problem !
  - der Fachbereich verantwortet  
*den Prozess, die Daten und deren Qualität, das Migrationsergebnis (muss damit leben und es abnehmen)*
  - nie darf die IT Bewertungen hinsichtlich  
*der Datenrelevanz (was muss migriert werden ?)*  
*der Datenqualität (ausreichend ?) treffen*
  - ➔ Hebammenfunktion der IT: bestmögliche Unterstützung im Migrationsprozess



## 2. goldene Migrationsregel



- perfekte Datenqualität: kein Unternehmen will sie, braucht sie oder wird dafür bezahlen
  - 80:20 Regel
  - der Fachbereich entscheidet, wie Qualitätsprobleme zu priorisieren sind oder ignoriert werden können (vgl. goldene Regel 1)
  - Gefahr des Overengineering:  
Migrationsanwendungen müssen nur  
*genau 1x richtig laufen und dabei  
im definiertem Migrationszeitfenster  
die vereinbarte Datenqualität liefern*



## 3. goldene Migrationsregel



- Was Du nicht zählen kannst, zählt nicht !
  - Migrationsausgangslage, -ziele und -fortschritte messbar machen
  - Transparenz des Migrationsfortschritts (= Teilprojektstatus) für den strukturierten Projektablauf eines Gesamtprojekts entscheidend
  - von qualitativen Aussagen zu quantitativen, z.B.
    - Anzahl zu migrierender Datensätze je Entität*
    - Anzahl aktuell migrierbarer Datensätze je Entität*
    - % Fertigstellungsgrad je Entität und gesamt*



# 4-Phasenmodell für Migrationsprojekte





# 4-Phasenmodell für Migrationsprojekte



typischer Startpunkt für viele Migrationsprojekte



## 4-Phasenmodell für Migrationsprojekte



- als separates Teilprojekt verstehen
- ersten Projektplan mit Projekt / Programm abstimmen
- Teamzusammenstellung und Kickoff
- erste Auflistung
  - Altdatenquellen
  - Altdaten-Owner
  - Schlüsselenitäten



## 4-Phasenmodell für Migrationsprojekte



- als separates Teilprojekt verstehen
- ersten Projektplan mit Projekt / Programm abstimmen
- Teamzusammenstellung und Kickoff
- erste Auflistung
  - Altdatenquellen
  - Altdaten-Owner
  - Schlüsselenitäten



# Wen brauchen wir ? Das Migrationsteam entsteht.



- konkrete Personen identifizieren
  - die Owner der Prozesse und der Altsysteme
  - Prozessexperten
  - technische Altdatenexperten
  - Programm- / Projektvertreter
  - Datenarchitekt des Unternehmens
  - Auditierungsexperten
  - Betriebsrat und Datenschutzbeauftragte
- Tipp: frühe Einbindung, denn deren Wissen und Entscheidungen sind erfolgskritisch

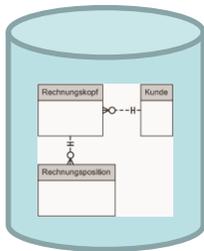




# Altdatenquellen, Owner und Schlüsselenitäten ermitteln



Altsysteme



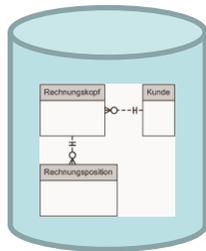
ORA DB von  
System 1  
Owner: TA-33



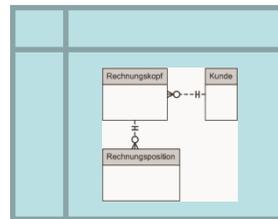
# Altdatenquellen, Owner und Schlüsselenitäten ermitteln



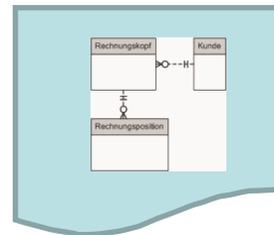
Altsysteme



ORA DB von  
System 1  
Owner: TA-33



Excel-Tabellen  
von System 2  
Owner: ???



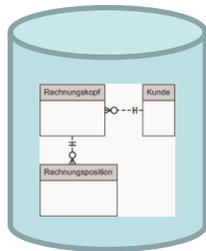
Flat Files von  
System 2  
Owner: ???



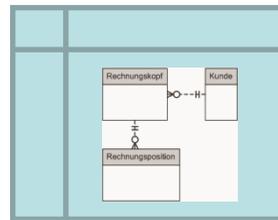
# Altdatenquellen, Owner und Schlüsselentitäten ermitteln



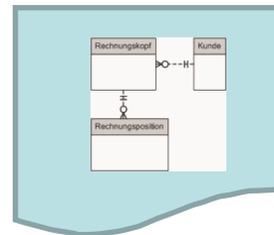
Altsysteme



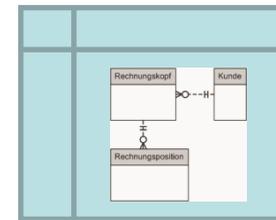
ORA DB von  
System 1  
Owner: TA-33



Excel-Tabellen  
von System 2  
Owner: ???



Flat Files von  
System 2  
Owner: ???



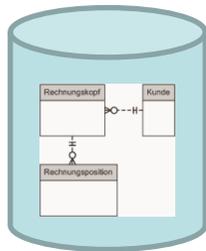
Access DB  
auf „privatem“  
System  
Owner: Fr. Meyer



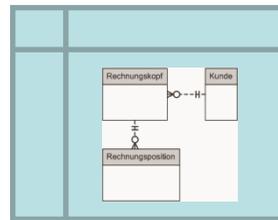
# Altdatenquellen, Owner und Schlüsselentitäten ermitteln



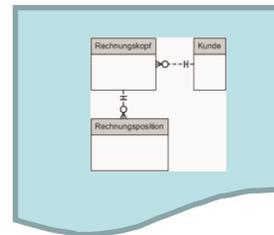
Altsysteme



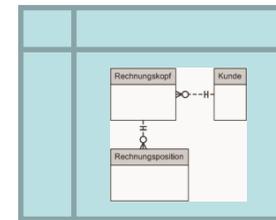
ORA DB von  
System 1  
Owner: TA-33



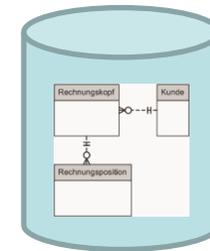
Excel-Tabellen  
von System 2  
Owner: ???



Flat Files von  
System 2  
Owner: ???



Access DB  
auf „privatem“  
System  
Owner: Fr. Meyer



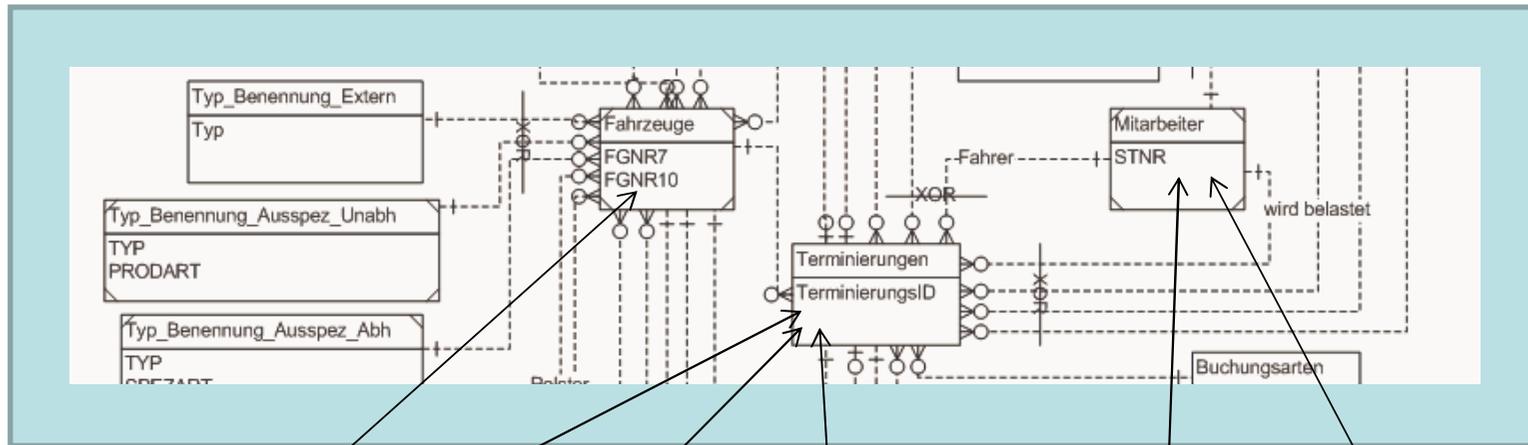
ORA DB von  
System 3  
Owner: PA-1



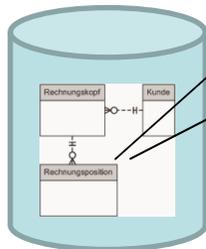
# Altdatenquellen, Owner und Schlüsselentitäten ermitteln



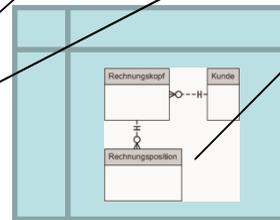
Neusystem



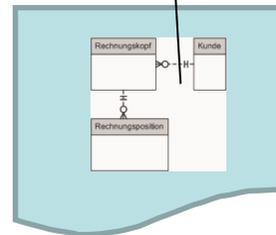
Altsysteme



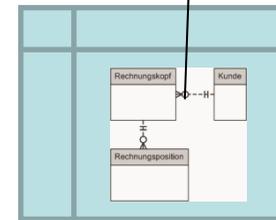
ORA DB von System 1  
Owner: TA-33



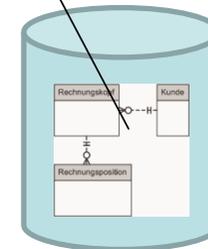
Excel-Tabellen von System 2  
Owner: ???



Flat Files von System 2  
Owner: ???



Access DB auf „privatem“ System  
Owner: Fr. Meyer



ORA DB von System 3  
Owner: PA-1



## 4-Phasenmodell für Migrationsprojekte



- Analyse  
Altdatenquellen
- Verbesserung  
Altdatenqualität
- Entwurf System  
Retirement Policy
- (Vor-)Auswahl  
Migrationstyp
- Abgleich Projekt-  
status mit -plan



# 4-Phasenmodell für Migrationsprojekte



- Analyse  
Altdatenquellen
- Verbesserung  
Altdatenqualität
- Entwurf System  
Retirement Policy
- (Vor-)Auswahl  
Migrationstyp
- Abgleich Projekt-  
status mit -plan



## Qualitätsdefizite entdecken / Anhebung der Altdatenqualität



- Workshops mit dem Migrationsteam
  - Erfassung der subjektiven Qualitätseinschätzung  
*dabei z.B. bekannte „Workarounds“ wie  
Feldmissbrauch aufdecken*
  - Ermittlung bisher unbekannter Probleme  
*z.B. Datenmodellierung inkonsistent  
gemeinsame Definition von Qualitätsprüfungen,  
z.B. range check, Typprüfung, ref. Integrität,  
Plausiprüfungen, Realitätscheck ...*



## Qualitätsdefizite entdecken / Anhebung der Altdatenqualität



- Workshops mit dem Migrationsteam
  - ➔ Definition einer Qualitätsmetrik incl. einer Zielvorgabe
  - bei Unterschreitung: Verbesserungsmaßnahmen definieren oder Absenkung der Zielvorgabe (mit Zustimmung Owner)
    - direkt im Altsystem*
    - falls unmöglich: auf temporären Datenbeständen*
  - ➔ aber: Regel 2 beachten (z.B. 80:20 Regel)



# System Retirement Policy



- Spezifikation und Plan darüber, wie ein System mit relevanten Altdatenbeständen außer Betrieb genommen wird
- Inhalt pro Altsystem mit Owner abgestimmt
  - was wird migriert, was nicht
  - Zugriffanforderungen auf nicht migrierte historischen Daten (auf welche, wer (Revision ?), wann, wie oft, mit welchem Vorlauf ?)
  - Zeitpunkt (x Wochen nach Go Live) und Art der Abschaltung (komplett, read-only, ...)



# Migrationstypen



- **Big Bang**
  - + einfach, kostengünstig
  - riskant
- **Parallelbetrieb**
  - + sicher, für Mission- or Safety-Critical Apps
  - schwierig, teuer
- **phasenhafte Migration**
  - + Lösung für kleinste Zeitfenster (z.B. nach Datenwichtigkeit oder nach Business Einheiten)
  - lange Projektdauer ist selber ein Risiko



## 4-Phasenmodell für Migrationsprojekte



- Fertigstellung Schlüsselenitäten mit Datenmodell und System Retirement Policy
- Zielsystemdatenmodell betrachten
- Matching der Schlüsselenitäten
- ETL Regeln definieren
- Timing sicherstellen
- Abgleich Projektstatus mit -plan



## 4-Phasenmodell für Migrationsprojekte



- Fertigstellung Schlüsselenitäten mit Datenmodell und System Retirement Policy
- Zielsystemdatenmodell betrachten
- Matching der Schlüsselenitäten
- ETL Regeln definieren
- Timing sicherstellen
- Abgleich Projektstatus mit -plan



- Workshops mit Migrationsteam
    - Vergleich Alt- und Neusystemdatenmodelle
    - Erarbeitung / Dokumentation der Mapping-Regeln
      - zunächst Entitätsebene, später Feldebene*
      - typisches Problem: Homonyme und Synonyme
- Datenstrukturänderungen*
- Datentransformationsregeln*
- Reihenfolge der Transformationen*
- Aufdeckung / Generierung von fehlenden Daten*



## Datenmapping [2/2]



- Ifd. Fortschrittsprüfung auf Entitätsebene durch Metrik und Zielgrößendefinition, z.B.
    - wie viele Sätze sind so nicht sinnvoll migrierbar ?*
    - wie viele Sätze fehlen aus Sicht Zielsystem ?*
  - sollte Zielgröße nicht erreicht werden: Im Team
    - Festlegung von Schritten zur Verbesserung der Datenqualität, oder*
    - bewußte Absenkung der Qualitätsanforderungen*
- ➔ vom owner abzuzeichnendes Ergebnis
- data gap*
- ETL Definitionen (Extract, Transform, Load)*



- Wahl des Migrationszeitfensters
  - harte Termine und Zeitfenster typisch: das Verstreichen kann bis zu 12 Monate Verzug kosten
  - gute Beziehungen zum Migrationsteam nutzen
- Migrationszeitfenster sicher einhalten
  - kritische Größe „Laufzeiten“
  - Datenvolumen und Anzahl Datensätze relevant
  - ETL-Tools i.d.R. langsam vgl. mit optimiertem Code, dafür einfacher in der Implementierung
  - Tests mit produktionsidentischer Plattform



## 4-Phasenmodell für Migrationsprojekte



- physisches Design
- Migrations-SW Build oder Tool-parametrisierungen
- Fallbackplan
- Migrationstestläufe incl. Verifikation Timing
- Migrations-durchführung
- lessons learned



# 4-Phasenmodell für Migrationsprojekte



- physisches Design
- Migrations-SW Build oder Tool-parametrisierungen
- Fallbackplan
- Migrationstestläufe incl. Verifikation Timing
- Migrations-durchführung
- lessons learned



# Ergebnisprüfung Migrationslauf (Test und echt)



- **Audit Trail / Migrationsprotokoll**
  - Protokoll aller Aktionen des ETL-Prozesses
  - dokumentiert Datensätze, die den Migrationsprozess wg. Qualitätsmängeln nicht überstehen
- **Kontrollsummen**
  - Summen über Werte oder Zahl übertragener Sätze
  - Vergleich Alt- / Neusystem = Audit Trail Delta ?
- **user acceptance tests**
  - vor go live durch den Prozessexperten
  - von „fühlt sich gut an“ bis „volles Testprogramm“



- Nehmen Sie Datenmigrationen ernst: Sie können schwierig oder gar gefährlich sein.
  - Goldene Regel 1: Eine Datenmigration und die erreichte Datenqualität ist ein fachliches und kein technisches Thema.
  - Goldene Regel 2: Kein Unternehmen bezahlt für perfekte Datenqualität.
  - Goldene Regel 3: Was Du nicht zählen kannst, zählt nicht.
  - Es gibt ein etabliertes 4-Phasen-Vorgehensmodell für Migrationen.
  - Lassen Sie Migrationen durch motivierte und erfahrene Kollegen oder alternativ durch Externe erledigen, wenn Sie Sorgen bzgl. des Aschenputteleffekts haben.
- Fazit: Auch wenn Datenmigrationen im Gutfall selten im Rampenlicht stehen, so sind sie bei Systemeinführungen von großer Kritikalität und verdienen entsprechende Beachtung durch die Gesamtprojektleitung.

