



## Aus alt mach neu!

### Data Cleansing & Migration von dezentralen Kundenstammdaten in ein Data-Warehouse-System



Um nicht nur heute, sondern auch in Zukunft, optimal für den Logistikmarkt gerüstet zu sein, investiert die Hellmann Worldwide Logistics GmbH & Co. KG derzeit sehr stark in ihre IT. Für die Sicherung einer mittel- und langfristig effektiven Kostenstruktur in den operativen fachlichen und technischen Bereichen, entschied man sich u. a. für eine Standardsoftware, die international ausgerollt werden sollte.

#### Die Problemstellung

##### Auf Anwendungskonsolidierung folgt Datenkonsolidierung

Einhergehend mit der Entscheidung für die Oracle E-Business Suite als zukünftige Standardsoftware wurde ein umfangreicher Transformationsplan ausgearbeitet, der auf mehrere Jahre ausgelegt war. Ein erster Meilenstein dieses Plans war der Aufbau und die unmittelbare Nutzung des zentralen Customer Data Hubs in der Oracle E-Business Suite. Für dieses Vorhaben musste die Datenpflege, die bis zu diesem Zeitpunkt noch in den einzelnen Legacy-Systemen stattfand, auf das neue zentrale System umgestellt und zuvor die Altdaten bereinigt und migriert werden. Da es sich bei der Übernahme weiterer Stamm- und Bewegungsdaten um einen wiederkehrenden Schritt handelte, der im Rahmen des geplanten Transformationsprozesses auch für nachfolgende Phasen geplant war, hieß das Ziel, einen weitestgehend generischen Data-Cleansing- & Migration-Prozess zu entwickeln und zu implementieren.

#### Der Kunden-Nutzen

##### Data Cleansing & Migration Prozess bringt Entlastung und wichtige Erkenntnisse

Viele Gründe sprachen in diesem Fall für eine umfangreiche Voruntersuchung und vorbereitende Maßnahmen – auch im Hinblick auf die weiteren Teilprojekte in diesem ersten Transformations-schritt zur Standardsoftware-einführung! Die Vorteile:

- Data Profiling lieferte wichtige Erkenntnisse für die Zieldatenstruktur: Der Input für die anderen Teilprojekte konnte hieraus generiert werden.
- Datenbereinigung konnte schon in den Legacy-Systemen durchgeführt werden: Die zu migrierende Datenmenge wurde deutlich reduziert.
- Generischer Prozess für noch anstehende Datenmigrationen und die Nutzung eines vielseitig einsetzbaren ETL-Tools: Synergieeffekte bei Know-how-Aufbau und -Nutzung sowie Personal- und Sachkosten.

#### Kundeninformationen



1871 von Carl Heinrich Hellmann gegründet hat sich das Unternehmen zu einem der großen internationalen Logistik-Anbieter entwickelt. Die Dienstleistungs-palette von Hellmann Worldwide Logistics umfasst dabei die klassischen Speditionsleistungen per LKW, Bahn, Luft- und Seefracht genauso wie ein umfangreiches Kurierdienstangebot, Warehousing und zahlreiche Spezial-Serviceleistungen.

**Nähere Informationen unter:** [www.hellmann.net](http://www.hellmann.net)

#### Projektziel

Bereinigung und Konsolidierung von Kundenstammdaten aus den Systemen für Landverkehre und Luft- und Seefracht sowie aus weiteren operativen Anwendungen wie z. B. CRM in den zentralen Data Hub einer neuen Standardsoftware.

#### Vorgehen

##### Generisch und inkrementell zum Ziel

Schritt 1: Planung und Konzeption

Im ersten Schritt wurden die zu migrierenden Systeme und die für diese erste Transformationsphase relevanten Daten bestimmt, Architektur- und Datenübertragungsszenarien aufgezeigt sowie mögliche Tools einem Auswahlprozess unterzogen und eine Data-Maintenance-Strategie festgelegt.

## Schritt 2: Quellsystemanalyse und Data Profiling

Darauf aufbauend wurden die im Scope befindlichen Legacy-Systeme nach zuvor definierten Kriterien untersucht sowie technische und fachliche Zusammenhänge und Abhängigkeiten aufgezeigt. Auf Basis dieser Erkenntnisse wurden daraufhin Daten aus den Legacy-Systemen extrahiert und mittels der Data-Quality-Option des Oracle Warehouse Builders analysiert.

## Schritt 3: Entwicklung der ETL-Strecke und Durchführung der Migration

Ebenfalls mit dem Oracle Warehouse Builder, welcher eigentlich eher als ein klassisches Extraktions-, Transformations- und Lade-(ETL-)Werkzeug im Bereich Business Intelligence und Data Warehouse bekannt ist, wurde die Migration entwickelt, ergänzt durch individuelle PL/SQL-Routinen als Schnittstelle zum Trading-Community-Architecture (TCA)-Modell des Customer Data Hubs.

# Lösungsszenario

## Optimales Kosten-/Nutzenverhältnis durch den Einsatz des Oracle Warehouse Builders und der Data-Quality-Option

Durch den Einsatz von OWB und der Data-Quality-Option wurde z. B. die SOUNDEX-Funktion von Oracle zur Erkennung von Doubletten genutzt, die semantische Korrektheit mit Referenzdatenbeständen abgeglichen und syntaktische Prüfungen sowohl mit regulären Ausdrücken als auch per Implementierung von Mappingtabellen realisiert.

### [E]xtraktion

Sämtliche Daten wurden zunächst 1:1 aus den Legacy-Systemen in eine sogenannte Staging Area eingeladen (analog zum Vorgehen bei einem Data Warehouse). Damit war eine erste Entkopplung erreicht und die darin befindlichen Daten konnten ohne weitere systemische Verbindung zu den Legacy-Systemen bearbeitet werden.

### [T]ransformation

Aus der Staging Area heraus wurden die Daten nach den Regeln, die zuvor aus dem Data Profiling ermittelt worden waren, in eine Working Area transformiert, welche sich bereits sehr nah am Zieldatenmodell orientiert.

### [L]aden

In diesem Schritt wurden die Daten aus der Working Area zunächst in sog. TCA-Interface-Tabellen und von dort nach Anwendung der über die Data-Librarian-Funktionalität direkt in der E-Business Suite implementierten Deduplication Rules in das produktive TCA-Zieldatenmodell des Customer Data Hubs geladen.

## Technik-Informationen

### Ausgangssituation:

- Zwei Kernsysteme für Landverkehre sowie Luft- und Seefracht (Individualsoftware)
- Ein Kernsystem für buchhalterische Aufgaben (Standardsoftware)
- Ca. 15 weitere Produktionssysteme (Individual- u. Standardsoftware)
- Länderspezifische Instanzen

### Die neue Lösung:

- Ein zentraler Customer Data Hub
- Einheitliche Logik zur Datenpflege
- Entsprechende Funktionen in den Legacy-Systemen deaktiviert
- Übergangsweise Rückführung der Daten aus dem Customer Data Hub über einen ESB zur Weiterverwendung in die Legacy-Systeme

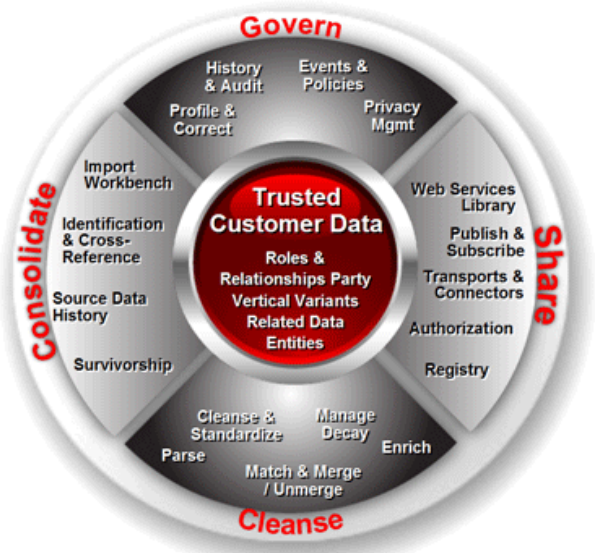


Abbildung: <http://www.oracle.com/master-data-management/cdh.html>

### Fazit:

#### Technologische Aspekte

- Einsatz der Oracle Data Quality Option für das Data Profiling
- Einsatz des Oracle Warehouse Builders für die Datenmigration
- Transformation der Legacy-System Daten in das TCA-Modell des Oracle Customer Data Hubs
- Einsatz der Data-Librarian-Funktionalitäten zur laufenden Sicherstellung der Datenqualität in der Oracle E-Business Suite

#### Fachliche Aspekte

- Vermeidung redundanter Datenerfassungen
- Regelwerke zur laufenden Datenqualitätssicherung
- Einführung von Data Stewards
- Generischer semiautomatischer Prozess



OPITZ CONSULTING

[www.opitz-consulting.com](http://www.opitz-consulting.com)