

Datenbankprojekt der FU Berlin

WS 2003/2004

Spezifikation für den Durchstich-Prototypen des ETL-Tools „ETLaien“

Stichtag/Deadline: 16.01.2004

1. Anforderung an den Durchstichprototypen:

Der Durchstich-Prototyp soll Daten aus einer Oracle-Datenbank (Source) auslesen und in eine andere Oracle-Datenbank (Target) überführen können.

Folgendes Szenario soll der Prototyp problemlos bewältigen können:

Auf der **Source-Datenbank** befinden sich mehrere Tabellen. Eine dieser Tabellen heißt Kunden_ADI_SUED, eine andere Kunden_ADI_NORD

Beide enthalten folgende drei Spalten:

Name	VARCHAR2 (30) primary key,
Age	NUMBER,
Colour_of_hair	VARCHAR2 (25).

Auf der **Target-Datenbank** befindet sich eine Tabelle (Kunden_ADI_Gesamt).

Sie enthält folgende zwei Spalten:

Name	VARCHAR2 (30) primary key,
Colour_of_hair	VARCHAR2 (25).

Mit Hilfe des ETL-Tools soll folgendes möglich sein:

- Verbindung zur Source-Datenbank aufbauen
- Verbindung zur Target-Datenbank aufbauen
- Tabelle „Kunden_ADI_SUED“ aus allen Tabellen der Source-Datenbank auswählen
- Spalten „Name“ und „Colour_of_hair“ für Übertragung selektieren
- Tabelle „Kunden_ADI_Gesamt“ als Ziel angeben
- Transfer aller Namen mit dazugehöriger Haarfarbe in Zieltabelle in die gleichnamigen Spalten
- Tabelle „Kunden_ADI_NORD“ aus allen Tabellen der Source-Datenbank auswählen
- Spalten „Name“ und „Colour_of_hair“ für Übertragung selektieren
- Tabelle „Kunden_ADI_Gesamt“ als Ziel angeben
- Transfer aller Namen mit dazugehöriger Haarfarbe in Zieltabelle in die gleichnamigen Spalten (hinter Daten von ADI_SUED)

2. Anforderungen im Einzelnen:

2.1 Anforderung an die GUI:

- Connection-Wizard (vollständig funktionsfähig)
 - o Wizard zum Verbindungsaufbau zur Source-/Target-Datenbank
- Transfer-Wizard (noch ohne spezielle Funktionen oder Join)
 - o 1. Schritt: Auswahl der Source-Tabelle(n) und Target-Tabelle
 - o 2. Schritt: Selektion der Source- und Target-Spalten
- Table-Wizard (nur wenn zeitlich möglich)
 - o Wizard zur Erstellung einer neuen Tabelle in der Target-DB

2.2 Anforderung an Import/Export:

- Verbindung zu Source-/Target-Datenbank aufbauen
- Alle Tabellen der Schemata auslesen
- Connection-Pooling

2.3 Anforderung an Core:

- Metadaten über notwendige Operationen/Funktionen, Tabellen und Spalten der Datenbank erstellen und in Repository verwalten

2.4 Anforderung an Worker:

- SQL-Statement erstellen
- Transfer von Source- nach Target-Datenbank
 - o „SELECT“ aus Source-DB
 - o Zwischenspeichern (in ResultSet)
 - o „INSERT“ in Target-DB

Alle implementierten Funktionen sind zu dokumentieren!!!