

## Heterogene Datenquellen zusammenführen - Load

REU/SDE/PIE



### Aufgabe 5 – Schritt L(oad)

Die im letzten Schritt müssen die zusammengeführten Daten nun in geeigneter Form extrahiert bzw. gespeichert werden. Üblicherweise geschieht dies in großen Data-Warehouse Umgebungen oder Datenbanken, die aktuell verwendet werden. Dadurch steigt die Komplexität des Schrittes „Load“ drastisch.

Wir beschränken uns jedoch hier auf unser Fallbeispiel, was den Load-Schritt somit im einfachsten Fall tatsächlich auf ein „export“ in eines der drei Datenformate beschränkt.

## Anforderungen/Kriterien zur Bewertung des Programms

Ihr Programm...

- ...speichert die resultierenden Daten Ihrer „Staging Area“ in allen drei Eingabeformaten, der Nutzer kann zwischen den drei Formaten wählen. Die Struktur muss **nicht** der Originalstruktur entsprechen (z.B. darf die resultierende XML Datei auch mehr/weniger Spalten enthalten).
- [optional] ...ermöglicht es, die resultierenden Daten Ihrer „Staging Area“ wieder in die Originaldateien zu integrieren (z.B. werden nur zusätzliche Zeilen/Spalten hinzugefügt, die Daten IN der Datei werden nicht adaptiert/gelöscht).
- ...protokolliert den gesamten Prozess in einer sinnvollen und nachvollziehbaren Form (z.B. in Form einer Log-Datei / Konsolenausgaben)
- ...ist vollständig kommentiert, insbesondere bei Verwendung von Bibliotheken.