

Bachelor-Thesis

Untersuchung des ETL-Tools Tensei-Data und dessen Einsatz im
Prozess der Datenmigration bei der SIV.AG

Eingereicht am: 21. Februar 2016

von: Martin Pohl
geboren am 13. April 1981
in Rostock

Aufgabenstellung

Einfügen der ausgegebenen Aufgabenstellung der Bachelor-Thesis. Der Titel der Arbeit wird bei deutschsprachigen Titeln in der englischen Fassung wiederholt.

Zusammenfassung

Hier sollte auf max. 1/2 bis 3/4 Seiten eine Zusammenfassung erstellt werden.

- Motivation, Einordnung, Umfeld und Abgrenzung der Arbeit
- wesentliche Schwerpunkte und Ergebnisse der Arbeit

Abstract

English Version.

Danksagung

An dieser Stelle möchte ich mich zunächst bei meiner Betreuerin an der Hochschule Wismar Prof. Dr.-Ing Antje Raab-Düsterhöft für ihre Unterstützung bedanken.

Diese Arbeit ist am Standort Roggentin der SIV.AG entstanden. Dafür möchte ich mich besonders bei meinem zweiten Betreuer, Detlef Herold, bedanken. Meinen Dank möchte ich auch allen anderen Mitarbeitern der SIV.AG aussprechen, die stets ein offenes Ohr für meine Fragen hatten.

Inhaltsverzeichnis

1	Einleitung	6
2	Theoretische Grundlagen	7
2.1	SIV.AG	7
2.2	kVASy®	7
2.3	Datenbankgrundlagen	8
2.3.1	Elemente einer relationalen Datenbank	8
2.3.2	Beziehungen und Normalformen	9
2.3.3	Schema	9
2.4	Datenintegration/Datenmigration	9
2.5	Transformationen	9
2.6	ETL-Prozess	9
2.6.1	Extraktion	10
2.6.2	Transformation	10
2.6.3	Laden	10
2.7	Qualitätskontrolle/Protokollierung	10
2.8	Kriterien für erfolgreiche Migrationen	10
2.8.1	Ziele einer Migration	11
2.8.2	Wirtschaftliche Aspekte	11
2.8.3	Qualitative Aspekte	11
2.8.4	Organisatorische Aspekte	11
2.8.5	Vorgehensweise/Migrationsplanung	11
2.8.6	Anforderung an die Datenmigration	11
3	Ist-Stand-Analyse	12
3.1	Migrationsprozess	12
3.2	Transformationen	12
3.3	Qualitätskontrolle/Protokollierung	12
3.4	Ablaufsteuerung	12
4	ETL-Tool Tensei-Data	13
4.1	Allgemeine Informationen	13
4.2	Migrationsprozess	13
4.3	Transformationen	13
4.4	Qualitätskontrolle/Protokollierung	13
4.5	Ablaufsteuerung	13
5	Implementierung und Test	14

6	Bewertung, Zusammenfassung und Ausblick	15
6.1	Bewertung	15
6.2	Ausblick	15
6.3	Zusammenfassung	15
	Literaturverzeichnis	16
	Abbildungsverzeichnis	17
	Tabellenverzeichnis	18
	Abkürzungsverzeichnis	19
A	Auflistung von Quellcode und ähnliches	20
	Selbstständigkeitserklärung	21

1 Einleitung

max. 2–3 Seiten

- Einführung in die Thematik und das wissenschaftlich-technische Umfeld
- Einordnung in das Wissenschaftsgebiet / tangierte Gebiete
- Analyse der Aufgabenstellung (Problemerkfassung)
- Realisierungsumfeld und Randbedingungen

2 Theoretische Grundlagen

In diesem Kapitel werden zunächst die Grundlagen und Begrifflichkeiten vorgestellt. Weiterhin werden die Kriterien für eine erfolgreiche Migrationen erläutert.

2.1 SIV.AG

Die SIV.AG ist ein Anbieter für ganzheitliche Lösungen im Bereich der deutschen und internationalen Energie- und Wasserwirtschaft. Das Kernstück des Unternehmens ist das Softwareprodukt kVASy[®], ein vollständig integriertes Enterprise Resource Planning (ERP)-System, welches insbesondere auf die Anforderungen und Prozesse der Energie- und Wasserwirtschaft ausgerichtet ist. Das Unternehmen wurde 1990 durch Jörg Sinnig als Software- und Beratungshaus gegründet und im Februar 2016 an die Harris Computer Corporation, einer Tochtergesellschaft der Constellation Software Inc. verkauft. Zum gegenwärtigen Zeitpunkt beschäftigt die SIV.AG mehr als 400 Mitarbeiter. Ihr Leistungsportfolio reicht von der Beratung und Analyse über die Implementierung und Bereitstellung der IT-Systeme bis hin zur Datenmigration, Schulung und Pflege. [SIV13]

2.2 kVASy[®]

Das zuvor bereits erwähnte kVASy[®] ist ein webbasiertes ERP-System und bildet das Aushängeschild der SIV.AG. Unter einem ERP-System ist eine integrierte unternehmensweite Anwendung zu verstehen, die zur Koordination wichtiger interner Prozesse eingesetzt wird [CPS10, Seite 482]. Zu der Produktfamilie von kVASy[®] gehören die Module Finance (Buchhaltung), Billing (Abrechnung) und EDM (Vertragsverwaltung), Technical Assets (Anlagenmanagement) und xRM (Kundenbeziehungsmanagement). Auf der Grundlage einer zentralen Datenbasis erfolgt die Kommunikation der Module und Funktionalitäten schnittstellenfrei [SIV13, Seite15]. Die Umsetzung dessen erfolgt durch das Datenbank-Management-System (DBMS) Database 11g Release 2 von der Firma Oracle. Für die Umsetzung der Applikationslogik wurde PL/SQL (Procedural Language/Structured Query Language) verwendet, eine proprietäre Programmiersprache der Firma Oracle [Wika].

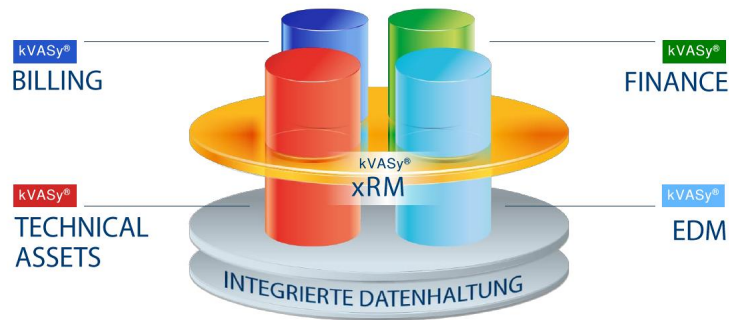


Bild 2.1: Produktfamilie von kVASy®[SIV16]

2.3 Datenbankgrundlagen

Indirekt wird bei Datenbanken von dem weit verbreiteten Relational Database Management System (RDBMS) ausgegangen. Dieses Datenbankmodell ist eins der Populärsten RDBMS [IT16]. Zu den Aufgaben eines Datenbanksystems gehört es die Daten möglichst effizient und dauerhaft zu speichern. Zudem soll es eine Vereinfachung der Datenverwaltung ermöglichen. Das relationale Datenmodell ist so konzipiert das es eine mengenorientierte Datenverarbeitung erlaubt. Dabei können die Daten aus mehreren Tabellen durch Join-Verknüpfungen in einer großen Tabelle abgebildet werden. Durch die Festlegung von Spalten und Where-Klauseln kann die Menge weiter bearbeitet und somit in der Höhe und Breite eingeschränkt werden. Im Gegensatz zu einem datensatzorientiertem Vorgehen wird die Tabelle, Zeile um Zeile abgerufen und nach bestimmten Kriterien geprüft. Nachdem die Kriterien erfüllt sind werden die Daten weiterverarbeitet. [KE11][Seite 71]

2.3.1 Elemente einer relationalen Datenbank

Zu den Elementen gehören Tabellen, Spalten und Zeilen (auch Tupel genannt). In der Regel existieren in einer Datenbank mehrere Tabellen. Die wiederum Attribute oder auch Felder beinhalten, die sogenannten Spalten. Eine Spalte besitzt einen Namen sowie einen Datentyp. Zu den drei wesentlichen Datentypen zählen Number (Zahl), Varchar (Text) und Date (Datum), diese enthalten wiederum noch weitere Varianten [KE11][Seite 112]. Die Bedeutung eines Datentyps ist zum einen für

ein Datenbanksystem und wichtig für ein Datenbanksystem für eine Optimierung des physikalischen Datenspeichers sowohl für Damit ist gewährleistet das nur ein bestimmter Datentyp aufgenommen werden kann.

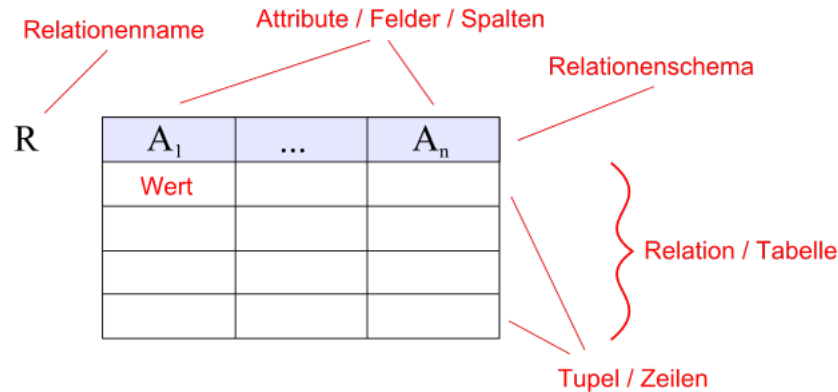


Bild 2.2: Begriffe einer relationaler Datenbanken [Wikib]

2.3.2 Beziehungen und Normalformen

2.3.3 Schema

2.4 Datenintegration/Datenmigration

Der Begriff Migration stammt von dem lateinischen Begriff migratio und bedeutet (Aus)wanderung, d. h. die Übernahme von Daten in ein anderes System, um sie dort wieder im vollem Ausmaß zu nutzen [Dud]. Die Datenmigration oder auch Datenübernahmen genannt, verläuft von einem Quellsystem hin zu einem Zielsystem. Dabei beinhaltet das Quellsystem, die zu migrierenden Daten und das Zielsystem die importierten Daten. In dieser Arbeit wird das in Kapitel 2.2 auf Seite 7 beschriebene kVASy[®], als Zielsystem verwendet.

2.5 Transformationen

2.6 ETL-Prozess

Die Abkürzung ETL setzt sich aus den drei Wörtern Extraktion, Transformation und Laden zusammen. Sie wird oft mit dem Begriff Data Warehouse in Zusammenhang gebracht. Dabei dient ETL dazu, unterschiedliche Bereiche Buchhaltung, Abrechnung, Vertragsverwaltung usw. in eine zentrale Datenbank (Data Warehouse) zusammenzuführen. Ziel ist es z. B. eine umfassende Sicht auf den Kunden zu

bekommen. Aufbauend darauf können komplexe Analysen entwickelt werden, um eine erfolgreiche und wohl überlegte Geschäftsentscheidung zu treffen. Bei der Abfolge eines ETL-Prozesses werden zunächst Daten aus einer Datenquelle extrahiert. Daraufhin werden diese mit Hilfe von Transformationsregeln homogenisiert. Schließlich werden im letzten Schritt die angereicherten oder bereinigten Daten in das Ziel geladen. Die Bandbreite für den Einsatz eines ETL-Prozesses erstreckt sich über das Datenqualitätsmanagement sowie der Replikation und Synchronisation von Daten bis hin zu der Migration von Daten aus Altsystemen. Letzteres ist für die Bachelorarbeit der wichtigste Aspekt und wird dahin gehend genauer untersucht. [Ros13]

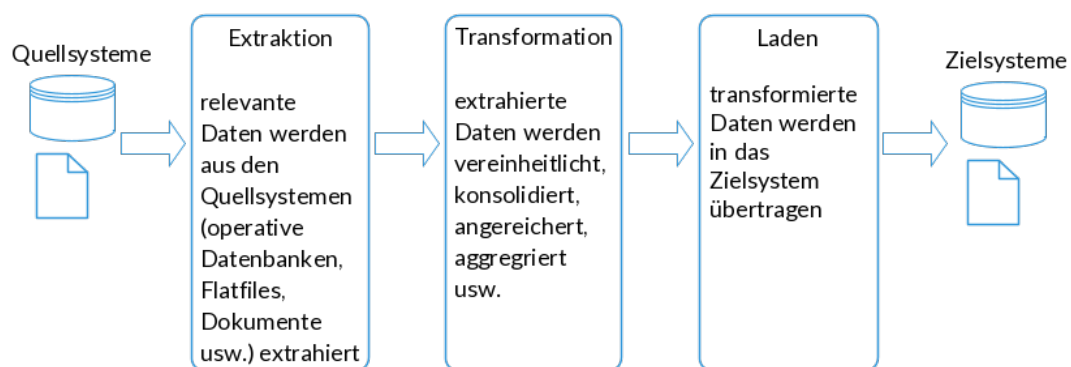


Bild 2.3: Allgemeiner ETL-Prozess (eig. Abb. nach [Ros13][Seite 38])

2.6.1 Extraktion

2.6.2 Transformation

2.6.3 Laden

2.7 Qualitätskontrolle/Protokollierung

2.8 Kriterien für erfolgreiche Migrationen

- Ziele einer Migration
- Vorgehensweise/Migrationsplanung
- Strategische Aspekte
- Rechtliche Aspekte
- Wirtschaftliche Aspekte

- Qualitative Aspekte
- Aspekte des Systembetriebs
- Organisatorische Aspekte
- Sicherheitsaspekte

2.8.1 Ziele einer Migration

Bevor es zur einer

- ein verbesserter Anwendernutzen
- das Herstellen eines rechtlich notwendigen Zustands
- die Behebung von Fehlern
- die Erweiterung des Funktionsumfangs
- eine verbesserte Integration in die vorhandenen Softwaresysteme
- eine verbesserte Interoperabilität
- eine Verringerung der laufenden Kosten
- die Erhöhung der Produktivität
- die bessere Nutzung vorhandener Ressourcen
- die Einhaltung strategischer Vorgaben

2.8.2 Wirtschaftliche Aspekte

2.8.3 Qualitative Aspekte

2.8.4 Organisatorische Aspekte

2.8.5 Vorgehensweise/Migrationsplanung

2.8.6 Anforderung an die Datenmigration

3 Ist-Stand-Analyse

3.1 Migrationsprozess

3.2 Transformationen

3.3 Qualitätskontrolle/Protokollierung

3.4 Ablaufsteuerung

4 ETL-Tool Tensei-Data

4.1 Allgemeine Informationen

4.2 Migrationsprozess

4.3 Transformationen

4.4 Qualitätskontrolle/Protokollierung

4.5 Ablaufsteuerung

5 Implementierung und Test

6 Bewertung, Zusammenfassung und Ausblick

6.1 Bewertung

6.2 Ausblick

6.3 Zusammenfassung

Literaturverzeichnis

- [CPS10] C.LAUDON, Kenneth ; P.LAUDON, Jane ; SCHODER, Detlef: *Wirtschaftsinformatik*. Pearson Studium, 2010
- [Dud] DUDEN: *Migration*. – <https://de.wikipedia.org/wiki/Migration> (Abrufdatum am 16.02.2016)
- [IT16] IT solid: *DB-Engines*. <http://db-engines.com/de/ranking>, 2016. – Abrufdatum: 17.02.2016
- [KE11] KEMPER, Alfons ; EICKLER, André: *Datenbanksysteme Eine Einführung*. Oldenbourg Verlag München, 2011
- [Ros13] ROSSAK, Ines: *Datenintegration*. Carl Hanser Verlag, 2013
- [SIV13] SIV.AG: *SIV.AG Imagebroschüre*. https://www.siv.de/tl_files/SIV/downloads/broschueren/SIV.AG_Imagebroschuere.pdf. https://www.siv.de/tl_files/SIV/downloads/broschueren/SIV.AG_Imagebroschuere.pdf. Version: 2013. – Abrufdatum: 01.02.2016
- [SIV16] SIV.AG: *SIV.AG Produktportfolio*. http://www.siv.de/tl_files/SIV/downloads/broschueren/SIV.AG_Produktportfolio.pdf. http://www.siv.de/tl_files/SIV/downloads/broschueren/SIV.AG_Produktportfolio.pdf. Version: 2016. – Abrufdatum: 21.02.2016
- [Wika] WIKIPEDIA: *PL/SQL* — *Wikipedia, Die freie Enzyklopädie*. – <https://de.wikipedia.org/wiki/PL/SQL> (Abrufdatum am 09.02.2016)
- [Wikb] WIKIPEDIA: *Relationale Datenbank* — *Wikipedia, Die freie Enzyklopädie*. – https://de.wikipedia.org/wiki/Relationale_Datenbank (Abrufdatum am 21.02.2016)

Abbildungsverzeichnis

2.1	Produktfamilie von kVASy [®] [SIV16]	8
2.2	Begriffe einer relationaler Datenbanken [Wikb]	9
2.3	Allgemeiner ETL-Prozess (eig. Abb. nach [Ros13][Seite 38])	10

Tabellenverzeichnis

Abkürzungsverzeichnis

ERP Enterprise Resource Planning
ETL Extraction, Transformation, Loading

A Auflistung von Quellcode und ähnliches

Selbstständigkeitserklärung

Hiermit erkläre ich, dass ich die hier vorliegende Arbeit selbstständig, ohne unerlaubte fremde Hilfe und nur unter Verwendung der aufgeführten Hilfsmittel angefertigt habe.

Ort, Datum

Unterschrift

Thesen
Bachelor-Thesis
Untersuchung des ETL-Tools Tensei-Data und dessen Einsatz im Prozess
der Datenmigration bei der SIV.AG

Eingereicht am: 21. Februar 2016

von: Martin Pohl
geboren am 13. April 1981
in Rostock

Betreuer: Prof. Dr.-Ing. Antje Raab-Düsterhöft
2. Prüfer: Betriebswirt (VWA) Detlef Herold

1. kurze Stichpunktartige Auflistung Diskussionswürdiger Punkte der Diplomarbeit
2. hier eine weitere These