

29. Juni 2009

## Bewerbung um den Hochschulpreis „David-Kopf“

Titel der im Sommersemester 2009 abgeschlossenen Diplomarbeit:

Qualitätskontrolle des ETL-Prozesses  
- Entwicklung eines qualitätsmetrik-basierten Ansatzes -

Verfasser: Christoph Metz

Erstbetreuer: Prof. Dr. Klaus Kruczynski

Zusammenfassung:

Datenqualität entscheidet über den Erfolg von Data-Warehouse-Systemen, die im Rahmen von Business Intelligence (BI) für die Unternehmenspraxis eine hervorragende Bedeutung erlangt haben. Der ETL-Prozess nimmt bei der Realisierung solcher Systeme etwa 70 bis 80 Prozent der Entwicklungsressourcen in Anspruch. Umso wichtiger ist es, geeignete Kontrollen der Prozessqualität durchzuführen. Die Unterschiede von dabei zu berücksichtigenden transaktionsorientierten (OLTP) und analyseorientierten (OLAP) Systemen erschweren diese Kontrollen. In Ermangelung umfassender Testszenarien für Data-Warehouse-Systeme muss zur Zeit auf Testkonzepte der allgemeinen Software-Entwicklung zurückgegriffen werden, die jedoch nur bedingt auf den ETL-Prozess übertragbar sind.

In der vorliegenden Arbeit werden die Unterschiede in Bezug auf das Testen der Datenqualität von OLTP- und OLAP-Systemen verdeutlicht. Darauf aufbauend, wird ein Ansatz entwickelt, der in vorhandene Testkonzepte unterstützend eingebracht werden kann. Der vorgeschlagene Ansatz basiert auf einer Folge von Quality Metrics, für die ein Qualitätsbegriff abgeleitet wird und für die exemplarisch Implementierungsmethoden aufgezeigt werden. Die Anwendung des Ansatzes ermöglicht eine kontinuierliche Messung und Analyse von Qualitätskennzahlen während des Prozessverlaufs. Dadurch lassen sich Fehler, die durch mangelnde Datenqualität verursacht werden, schnell einem bestimmten Prozessschritt zuordnen und beseitigen.

Der entwickelte Ansatz der Quality Metrics ist eine anwendungsbereite Bereicherung der BI-Theorie und -Praxis. Damit wird ein wirksamer Beitrag geleistet, die IT-Unterstützung für unternehmerische Entscheidungsprozesse fundiert zu verbessern.