## Wie misst man Daten- und Informationsqualität in IT-Systemen?

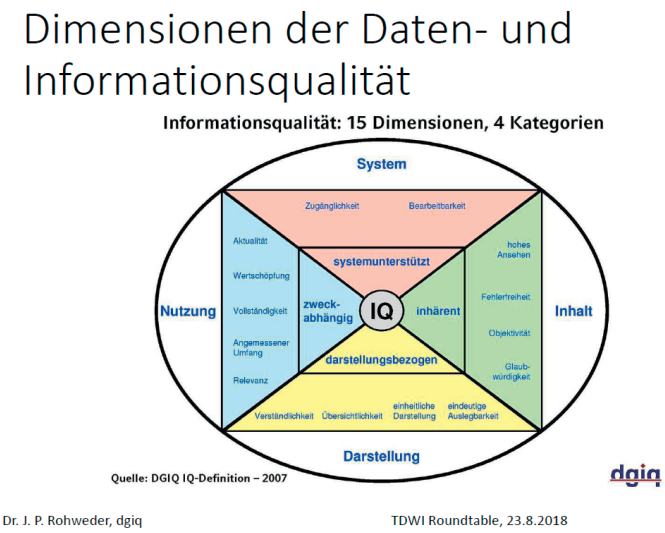
Der zweite Hamburger TDWI Roundtable in 2018 fand am 23. August im Gästehaus der Universität Hamburg statt. Bei strahlendem Wetter und viel zu hohen Temperaturen kamen mehr als 70 Teilnehmer aus Wirtschaft und Wissenschaft zu diesem Event, was sicher am interessanten Thema Datenqualität lag.

Wie immer wurde aufgrund der vielen Neuzugänge und Erstbesucher zuerst der Verein ausführlich von Georg Franzke, einem der Vorsitzenden des TDWI Roundtables Hamburg, vorgestellt. Er stellte die Strukturen und die vielen Informationen dar, die der Verein seinen Mitgliedern zur Verfügung stellt und somit wurde auch klar aufgezeigt, warum es sich für jeden lohnt, Mitglied des TDWI Vereins zu werden. Auch erläuterte er die vielfältigen Möglichkeiten, wie man sich im TDWI e.V. oder auch im Hamburger Roundtable engagieren kann.

Nach der Vorstellung des Vereins leitete er in das Thema des Abends durch die Vorstellung einiger Bücher zum Thema Datenqualität über. Vorgestellt wurde dabei in aller Kürze im Einzelnen die folgenden Bücher

* „Datenqualität für Entscheider in Unternehmen: Ein Dialog zwischen einem Unternehmenslenker und einem DQ-Experten“ von Guilherme Morbey
* „Datenqualität erfolgreich steuern: Praxislösungen für Business-Intelligence-Projekte (Edition TDWI)“ von Detlef Apel und Wolfgang Behme
* „Informationsqualität bewerten: Grundlagen, Methoden, Praxisbeispiele“ von Andrea Piro (Hrsg.)

Das letzte vorgestellte Buch war der Übergang zum Vortragenden des Abends, da Herr Dr. Rohweder Mitautor des Buches ist. Im folgenden Vortrag ging es vornehmlich um das Messen von Datenqualität. Dabei erwähnte der Sprecher am Anfang, dass man aufgrund von Untersuchungen auf eine Anzahl von 15 Dimensionen gekommen ist, mit denen man alle möglichen Probleme von Datenqualität abdecken kann. Eine Übersicht dieser Dimensionen zeigt die folgende Grafik, die aus dem Vortrag übernommen wurde:




Dr. Rohweder erklärte dann exemplarisch anhand von zwei Dimensionen das Vorgehen und den Aufbau des Buches, beim Buch hatte man sich an eine klar definierte Struktur gehalten.

Der Aufbau erfolgt dabei immer in den Schritten

- Definition der Dimension
- Abgrenzung zu anderen Dimensionen
- Messung der Dimension (oder alternativ Ermittlung der Dimension, wenn eine Messung nicht wirtschaftlich erfolgen kann).
- Anwendungsbeispiele
- Fazit & Interpretation

Für den Bereich der Fehlerfreiheit ergibt sich dadurch z.B. folgende Darstellung:

**Am Beispiel *Fehlerfreiheit:***

1. Definition: „Informationen sind fehlerfrei, wenn sie mit der Realität übereinstimmen“
2. Abgrenzung: ***Fehlerfreiheit*** ist nicht gleichbedeutend mit Datenqualität allgemein
3. Messmethode(n): objektive und subjektive Bewertungsmethoden
4. - Kennzahlen für die ***Fehlerfreiheit*** von Datenbankobjekten
   - Bewertung der ***Fehlerfreiheit*** aus Anwendersicht

Wichtig sind die Abgrenzungen zu anderen Dimensionen und auch der Hinweis, dass diese Abgrenzung nicht immer leicht ist, aber lohnenswert, um die Messmethoden richtig einzusetzen.

Neben der Dimension Fehlerfreiheit diskutierten die Anwesenden dann die Dimension Aktualität. Hier wies Dr. Rohweder wies darauf hin, dass Datenqualität einer zeitlichen Veränderung unterliegt und daher diese Änderung mit dem Verfallsratenmodell verglichen werden muss (was einen sehr interessanten Ansatz darstellt!).

Nicht verwunderlich war es, dass eine Vielzahl an Fragen zu dem Thema kamen. Es wurde versucht diese gemeinschaftlich zu beantworten. Schwierig in kurzer Zeit zu beantworten waren dabei Fragen, die eher allgemein gehalten waren, z.B. wie man ein DQ-Projekt am besten starten sollte (und dann das Management hinter sich bringen, so dass dieses die Finanzierung zugesagt). Es wurde der Hinweis gegeben, ein DQ-Projekt mit einer Analyse (Profiling) von wichtigen Daten zu starten, um einen ersten Eindruck von der Qualität zu erhalten und evtl. auch einen Business Case zu rechnen. Einfacher ist es in den Fällen, wo Unternehmen regulatorischen Anforderungen unterliegen (z.B. BCBS239, Solvency II, etc.), da hier eine Messung der Datenqualität von Behördenseite gefordert ist und somit auch der Task des DQ-Projektes der Ausschreibung und des Kundenauftrages ist. Damit ist die Finanzierung automatisch gesichert.

Wie immer wurde dann im Anschluss bei Bier und Brötchen das Thema weiter in kleiner Runde vertieft und sich auch sonst über aktuelle Projekte ausgetauscht. Und trotz der hohen Temperaturen blieben viele Gäste bis zum Ende - 21:30Uhr - wie immer. Dies zeigt, dass auch der Community-Charakter der Veranstaltung einen hohen Stellenwert hat.

**Die Vorsitzenden des Roundtables freuen sich schon auf die nächste Veranstaltung am 07.11.2018, wenn es um das Thema Blockchain geht.**

*Georg Franzke*
*unterstützt vom Vorstandsteam*