

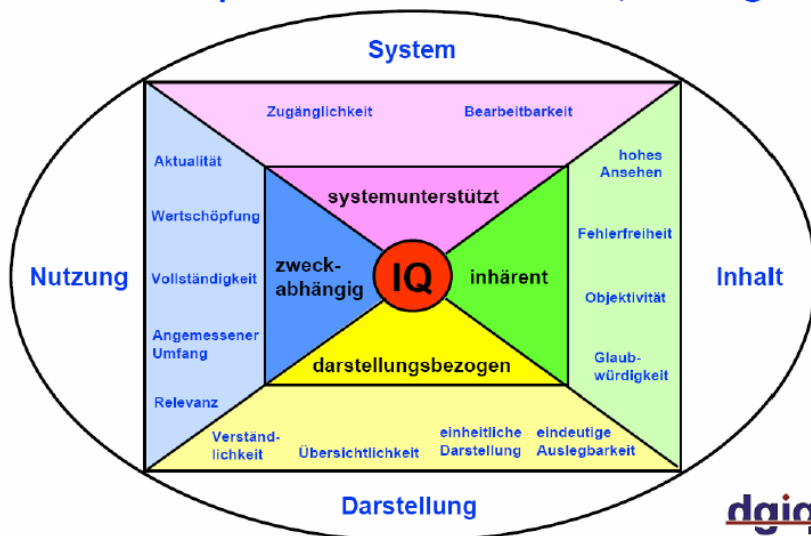
## Datenqualität: Der Schlüssel zum Vertriebs- und Marketingerfolg

„Shit in, shit out“ - dieser bekannte Spruch aus den Anfangszeiten der elektronischen Datenverarbeitung hat seine Aktualität noch längst nicht verloren. Im Gegenteil: Man staunt oft über die Naivität von Unternehmen, die meinen, mit dem Kauf eines IT-Systems und dessen Installation sei es schon betriebsbereit und stiftet den erwarteten Nutzen. Ein IT-System ist (immer noch) ein System, das Daten verarbeitet und daraus Informationen für die Anwender generiert. Wenn aber diese Daten nicht, unvollständig oder unkorrekt vorliegen, dann kann auch das beste IT-System keine wertvollen Informationen daraus ableiten. Übrigens: Auch wenn auf IT-Messen viele Anbieter nicht müde werden, dem Interessenten dennoch vorzugaukeln, dass - quasi auf magische Weise - mit ihrem System aus Datenmüll hochwertige Informationen erzeugt werden können.

### Zum Begriff der Datenqualität

Als Ausgangspunkt für eine Definition des Begriffs Datenqualität werden üblicherweise (siehe <http://www.dgiq.de>) die Arbeiten von R. Wang genommen (siehe <http://mitiq.mit.edu/publications.htm>); Wang kommt zu folgenden Kriterien (Quelle: <http://www.dgiq.de>)

### Informationsqualität: 15 Dimensionen, 4 Kategorien



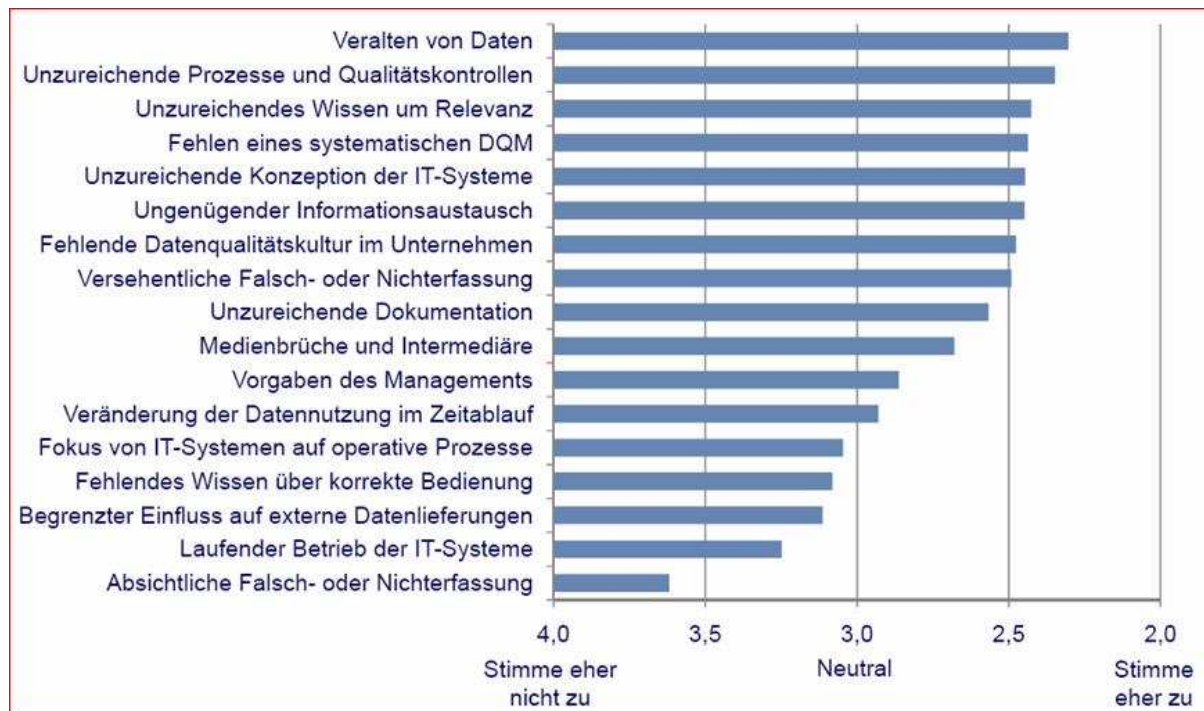
Während system- und darstellungsbezogene Aspekte zumindest zum Teil auf das verarbeitende IT-System verlagert werden können, müssen die Daten selbst insbesondere folgende Kriterien erfüllen:

- **Fehlerfreiheit:** Die Daten müssen mit der Realität übereinstimmen. Namen müssen richtig geschrieben sein, Größenangaben müssen im richtigen Maßstab angegeben sein, Artikelnummern müssen dem vorgegebenen Schema entsprechen, bei Firmenangaben muss die Rechtsformbezeichnung korrekt sein usw.
- **Aktualität:** Einmal (richtig) erfasste Daten müssen an eine veränderte Realität permanent angepasst werden. Über Datenobjekte (z.B. Kunden, Produkte, Lieferanten, Mitarbeiter etc.) muss der aktuellste Zustand gespeichert sein. Bei einer Heirat muss der Kundenname geändert werden, den Bilanzangaben über Firmenkunden muss der letzt verfügbare Jahresabschluss zugrunde liegen, usw.
- **Vollständigkeit:** Alle für einen Anwender relevanten Angaben über Datenobjekte müssen erfasst sein. Zusätzliche Telefonnummern von Kunden müssen erfasst werden, alle wichtigen Ansprechpartner mit vollständigen Kontaktangaben müssen vorliegen usw.

- **Angemessener Umfang:** Dabei sollten nicht alle verfügbaren, sondern nur alle relevanten Daten erfasst sein. Die Haarfarbe als Attribut mag für ein Friseurgeschäft eine wichtige Kundeninformation sein; für Maschinenbau-Unternehmen normalerweise nicht. Insbesondere sollten nur Daten in der Datenbasis vorhanden sein, die nicht bereits aus anderen gespeicherten Daten hervorgehen (**Redundanzfreiheit**): Es sollte nicht das Alter, sondern das Geburtsdatum gespeichert werden, da daraus das jeweilige Alter errechnet werden kann. Ein Attribut „Geschlecht“ genügt; alle Anredeformen in Anschreiben können daraus automatisch generiert werden und müssen nicht gesondert abgespeichert werden.

### Das Problem

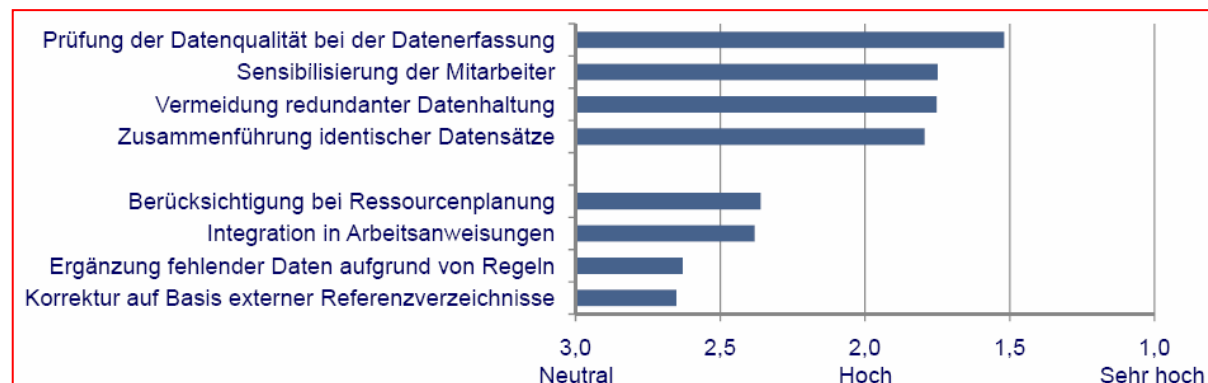
Dass in der betrieblichen Praxis gegen diese Kriterien massiv verstoßen wird, zeigt eine Umfrage von Wilde zum „Stand des Datenqualitätsmanagements im CRM“, präsentiert auf dem 13. CRM-Symposium am 23.04.2009 (<http://imari.dhbw-heidenheim.de/url/irat1106-1>); die Ursachen für Datenqualitätsprobleme werden vor allem gesehen in:



Versehen und Eingabefehler sind also nicht das Hauptproblem: Vor allem das **Vertrauen auf die Qualität bereits erhobener und qualitätsgeprüfter Daten** stellt das Hauptproblem dar. Da ja die Daten bereits einmal kontrolliert wurden, wird keine Veranlassung dazu gesehen, sie immer wieder und systematisch einer Qualitätskontrolle zu unterziehen. Die Daten veralten; eine fehlende „Datenqualitätskultur“ und fehlende organisatorische Voraussetzungen (z.B. ein Datenqualitätsbeauftragter) führen dazu, dass Qualitätsmängel bewusst oder unbewusst übersehen werden.

### Lösungen

Daraus lassen sich Konsequenzen für Maßnahmen ableiten; Wilde (s. oben) sieht hierbei in erster Linie organisatorische und führungsbezogene Maßnahmen:



Solche Maßnahmen können grundsätzlich unterteilt werden in

- **organisatorische Maßnahmen:** Datenqualität kann verankert werden
  - in der **Aufbauorganisation** durch
    - die Schaffung einer (Stabs-)Stelle „**Datenqualitätsbeauftragter**“, die - direkt der Unternehmensleitung unterstellt - die Datenqualität laufend kontrolliert (Qualitätsmonitoring) und durch das Einleiten, Durchführen und Überwachen von Datenqualitätssicherungs-Maßnahmen aktiv für Datenqualität sorgt (<http://imari.dhbw-heidenheim.de/url/lrat1106-2>).
    - die Verankerung des Unternehmensziels „Datenqualität“ als Daueraufgabe in den **Stellenbeschreibungen** aller Stellen, die mit der Erfassung, Verwaltung und Verwendung von Daten zu tun haben. So kann beispielsweise die Hotline damit beauftragt werden, bei jedem Kundenkontakt die wichtigsten Kundendaten abzufragen und den Datenbestand gegebenenfalls zu aktualisieren bzw. bei fehlenden Daten zu ergänzen.
  - in der **Arbeitsorganisation** durch ein **Anreizsystem**, das Mitarbeiter dafür be-/entlohnt, wenn sie für Datenqualität sorgen. Dies setzt entsprechende Messkriterien voraus, die dies feststellen. So könnte sich beispielsweise die Vergütung oder Provision daran ausrichten, wie oft Adressfehler in den von diesem Mitarbeiter betreuten („Owner“) Datensätzen auftauchen.
  - in der **Ablauforganisation** durch die Festlegung von **Datenqualitätssicherungsprozessen** (Wer hat wann mit welchen Maßnahmen für Datenqualität zu sorgen?). So könnte beispielsweise festgelegt werden, dass immer am Monatsanfang bestimmte Prüfroutinen (z.B. Dubletten-Check) ablaufen müssen.
- **technische Maßnahmen:** Das verwendete IT-System (z.B. das CRM-System) kann entsprechende Funktionalitäten zur Datenqualitätssicherung enthalten:
  - Bereits bei der Erfassung der Daten helfen **Erfassungsfunktionalitäten**, Eingabefehler zu vermeiden:
    - **Erfassungsautomatismen** wie beispielsweise das automatische Generieren von Adressdatensätzen aus E-Mail-Signaturtexten oder aus Telefonverzeichnissen (über die Telefonnummer) sparen nicht nur Eingabeaufwand, sondern reduzieren auch Eingabefehler.
    - **Importfunktionen** erlauben, Mengen von Datensätzen komplett zu erfassen. So kann man aus Datenbeständen über Exportfunktionen (z.B. aus Sozialen Netzwerken wie <http://www.xing.com/de/>) Importdateien für das eigene System erstellen und kann auf manuelles Eintippen (und damit Eingabefehler) verzichten.
    - **Schnittstellen** zu internen Datenbeständen (ERP- bzw. Warenwirtschaftssysteme, Office-Systeme wie beispielsweise MS Outlook) sowie zu externen Datenbeständen (z.B. Adressdatenbanken wie beispielsweise <http://shop.schober.com> statistische Datenbanken wie <http://www.destatis.de> oder Firmeninformationsdatenbanken wie <http://www.genios.de>) erlauben den Zugriff auf jeweils **aktuellste** Daten.
  - Bei bereits erfassten Daten helfen **Prüfroutinen**, die Datenqualität zu prüfen bzw. zu gewährleisten:
    - **Dublettenmanagement**-Funktionen sorgen dafür, dass Datensätze erkannt werden, die voraussichtlich (Sind „Meyer“ und „Mayer“ unterschiedliche Personen oder nur von verschiedenen Eingebenen unterschiedlich erfasste Namen derselben Person?) doppelt vorhanden sind ([http://www.crmmanager.de/magazin/artikel\\_1493\\_adressen\\_adressqualitaet.html](http://www.crmmanager.de/magazin/artikel_1493_adressen_adressqualitaet.html)). Die Ergebnisse werden dann von einem Sachbearbeiter beurteilt und in Korrekturmaßnahmen umgesetzt.
    - **Semantische Integritäts-Trigger** sind automatisch ablaufende Programme, die bei der Dateneingabe oder turnusmäßig danach die Daten auf bestimmte **Plausibilitäten** überprüfen. So kann bei jeder Ortsangabe automatisch geprüft werden, ob die zugeordnete Postleitzahl dazu passt (oder umgekehrt). Dies erfolgt durch Vergleich mit entsprechenden Adressdatenbanken.
- **Motivationsmaßnahmen:** Durch laufende Sensibilisierung der Mitarbeiter sowie durch Schulungsmaßnahmen (Was kann passieren? Welche technischen Datenqualitätssicherungs-Funktionen sind vorhanden? Wie funktionieren diese? Wer ist Ansprechpartner bei Problemen?) kann ein Bewusstsein über die Bedeutung der Datenqualitätssicherung („**Datenqualitäts-Kultur**“)

erzeugt werden. Dies funktioniert aber nur, wenn diese Maßnahmen über die Führungsebenen nach unten transportiert und vor allem vorgelebt werden. So sollten beispielsweise Außendienstmitarbeiter immer wieder dazu angehalten werden, ihr Wissen über ihre Kunden im System zu erfassen und die Daten der Kunden zu prüfen - soweit es datenschutzrechtliche Bestimmungen erlauben.

- **Externe Dienstleister:** Datenqualitätssicherung kann auch - in beschränktem Maße - auf darauf spezialisierte Dienstleister ausgelagert werden (<http://imari.dhbw-heidenheim.de/url/lrat1106-3>).

Eines sollte einem Auftraggeber aber immer bewusst sein: An Dienstleister und Software-Anbieter kann man nur Arbeiten auslagern, nicht aber die Verantwortung für die Datenqualität. Diese verbleibt beim Unternehmer.

#### **CEBIS hilft weiter**

**Unternehmen, die Informations- und Beratungsbedarf zu Fragen der Datenqualität oder allgemein zu Chancen, aber auch Risiken von IT und Internet haben, können sich an CEBIS wenden. In Veranstaltungen (wie beispielsweise am 07.07.2011 zum Thema Soziale Medien (Facebook, YouTube, XING,...) und in Einzelberatungen können Unternehmen Hilfestellung durch kompetente Experten erhalten. Informieren Sie sich auf der CEBIS-Website über die entsprechenden Termine und melden Sie sich möglichst frühzeitig an.**

Quelle und Copyright: Internetauftritt des Landkreises Neu-Ulm, <http://www.landkreis.neu-ulm.de>

Tipp des Monats **Juni 2011**