

Workshop

Datenqualität – Ein zumeist unterschätzter Erfolgsfaktor

Digital erfasste und gespeicherte Daten sind Grundlage vieler Prozesse in modernen Unternehmen. Nicht nur hochentwickelte Anwendungen, wie beispielsweise die Entscheidungsunterstützung im Rahmen des Business Intelligence, basieren mittlerweile auf einer breiten Datenbasis. In praktischen Anwendungen zeigt sich jedoch allzuoft die Unzulänglichkeit der gespeicherten Daten. In einer Studie sieht die Meta Group als Grund für das Scheitern von IT-Projekten in mehr als 35% der Fälle Mängel in der Datenqualität. Die durch Datenbestände von ungenügender Qualität entstehenden Mehrkosten schätzt die Meta Group allein in den USA auf zwei bis vier Milliarden US-Dollar pro Jahr. Von noch weit höheren Kosten spricht das Data Warehousing Institute. Vor dem Hintergrund dieser offensichtlichen Probleme und den damit verbundenen hohen Kosten ist der bisherige Stand der Forschung zum Thema Datenqualität als noch vergleichsweise bescheiden anzusehen.

Der Workshop stellt sich die Aufgabe, Wissenschaftler und Praktiker zusammenzubringen, um beide Gruppen für das Thema Datenqualität zu sensibilisieren und neue Anstöße zur Lösung von Datenqualitätsproblemen zu geben. Ziel ist es zum einen, von Seiten der Praxis konkrete Problemstellungen zu formulieren. Zum anderen soll anhand präsentierter praktischer Probleme Wissenschaftlern die Möglichkeit gegeben werden, gezielt Lösungsansätze aus anderen Gebieten zu übertragen oder neu zu entwickeln. Konkret stellen sich unter anderem die folgenden Fragen:

- Was ist Datenqualität? Wie lässt sich Datenqualität beschreiben, messen, operationalisieren? Welche Qualitätsdimensionen sind in der Praxis als relevant anzusehen?
- In welchen Anwendungen zeigt sich Datenqualität bereits heute als problematisch? Welche Anwendungsbereiche sind zu unterscheiden, und welche Qualitätsprobleme sind für bestimmte Typen von Daten charakteristisch, z.B. postalische Adressdaten, Bauzustandsdaten, Controllingdaten, ...?
- Welche strategische Bedeutung hat Datenqualität und wie lässt sich Management Attention erreichen? Welche organisatorische Einbindung ist anzustreben? Welcher Nutzen ist von guter Datenqualität zu erwarten? Welche Kosten lassen sich schlechter Datenqualität zurechnen?
- Welche Ansätze zur Datenqualitätsverbesserung sind denkbar und vielversprechend? Wann sind proaktive und wann reaktive Maßnahmen geeigneter? Sind Methoden aus den Bereichen Qualitätssicherung, Prozesssteuerung, Statistik und Data Mining im Bereich Datenqualität einsetzbar bzw. auf Datenqualitätsprobleme übertragbar?

Unter diesen thematischen Vorgaben war ausdrücklich auch zur Einreichung von Beiträgen zu praktischen Erfahrungen mit Datenqualitätsproblemen aufgefordert, unabhängig davon, ob bereits konkrete Lösungen gefunden werden konnten oder nicht. Aufgrund der hohen Resonanz auf den “Call for Papers” konnten lediglich 55% der Einreichungen zur Veröffentlichung berücksichtigt werden.

Ganz besonders bedanken möchte ich mich bei den zahlreichen Einsendern, beim Programmkomitee und den Organisatoren der GI-Jahrestagung 2004. Am Programmkomitee wirkten folgende Personen mit: M. Gebauer (WestLB AG), M. Grabert (DaimlerChrysler AG), U. Grimmer (DaimlerChrysler AG), U. Güntzer (Uni Tübingen), M. Helfert (Dublin City University), H. Hinrichs (FH Lübeck), M. Mielke (Deutsche Bahn AG), G. Nakhaeizadeh (DaimlerChrysler AG), F. Naumann (Humboldt-Universität Berlin) und U. Windheuser (WestLB AG)

Jochen Hipp (DaimlerChrysler AG)