

5 Millionen Verbindungsdatensätze

Dipl.Ing.Helmut Kropp

"Anbei erhalten Sie die Akten samt CD-ROM mit dem Auftrag, ein schriftliches Gutachten zur Richtigkeit der Abrechnung zu erstellen..."

Groß die Überraschung: auf der CD befinden sich 5 Millionen Datensätze, die ein Verbindungsnetzbetreiber einem anderen als Abrechnung übersandt hat. Diese Abrechnung ist strittig.

Da geht also nichts mit Handarbeit oder MS Excel. Ich denke da gar schon an Oracle und deren Datenbanken. Sicherheitshalber frage ich unseren SV-EDV-Experten, der empfiehlt mir Microsoft Access. Auf meine Bedenken hin meint der: "Also so schlecht ist Microsoft auch wieder nicht!"

Wenn schon, denn schon: kurzfristig buche ich einen Express-Kurs bei einem nahegelegenen Computerhaus für MS Access. Nach dessen Abschluß habe ich aber den Eindruck, dass mir doch noch viel an Know-how für eine erfolgreiche Arbeit fehlt. Also wird Access gekauft, ohne Probleme installiert und selbst getestet.

Dabei stellt sich heraus, dass das Abfragen der Datenbank gar nicht so das Problem ist, das macht sogar Spass. Weniger Spass macht aber das Einlesen der 5 Mio.Datensätze von der CD. Diese sind zwar in Klartext (ASCII), aber Access mag diese Zahlen nicht. Bis ich auf die Idee komme, die Felder der Datenbank als "Text" zu definieren: sofort ist die Datenbank da (Korrektur: mit dem etwas älteren Dell-PC mit WIN98 dauert es etwa 9-12 Minuten).

Nun kann es losgehen, Access stört sich gar nicht an der Felddefinition als Text, es rechnet mit diesen Feldern problemlos.

Aus den Akten gehen Tag- und Nachttarife etc. hervor, das ist rasch geprüft, Dauer und Kosten der Verbindung sind schnell kontrolliert. Einige Fehler sind schon zu erkennen, da wurde ein höherer Nacht-Tarif als vereinbart verwendet.

Große Überraschung dann bei den selbst gewählten Plausibilitäts-Checks. Nach einem Durchlauf durch die gesamte Datenbank fehlen immer 400 Datensätze. Unter 5 Mio.Datensätze die 400 fehlenden zu suchen, ist aber extrem zeitaufwändig.

Die Idee (nach dem Lesen aller möglichen Abfrageroutinen im Handbuch): Abfrage auf Feldinhalt Null!

Und tatsächlich: da sind schon die vierhundert fehlenden Datensätze. Sie haben alle ein Feld "Verbindungsdauer" mit der Größe Null, das Feld "Kosten" hat aber Inhalt. Das ist nun

logisch nicht zu erklären, denn bei einer Verbindungsdauer von Null müssen auch die Verbindungskosten Null sein, ganz gleich, wie hoch der Tarif ist. Und eigentlich sollte in so einem Fall gar kein Verbindungsdatensatz erstellt werden.

Nun kann man das Testen beenden und das Gutachten schreiben:

"Aufgrund der übergebenen Datensätze läßt sich die Abrechnung nicht nachvollziehen".

Schlußbemerkung: Etwa 3 Monate nach Abgabe des Gutachtens ruft der Rechtsanwalt an: "Also das ist ja ein starkes Stück, zu behaupten, dass unsere Abrechnung nicht nachvollziehbar ist!"

Antwort: siehe oben, Felder mit Inhalt Null. Wo bleibt die Logik?

Rechtsanwalt: "Also ja, ähem...Ich hätte da eine Bitte: Ich finde unsere CD nicht mehr, können Sie mir die CD schicken?"

Antwort: gerne, aber nur eine Kopie und über das Gericht.

Zusammenfassung: Nichts mehr davon gehört, Rechnung ohne Abzug bezahlt.

hk
09-2009