

# Datengetriebene kontra toolgetriebene Migration

VON JÜRIG MATHYS

**Die Datenmigration ist ein schwieriges Vorhaben. Meistens als ein Teil der Entwicklung und Einführung eines Neu-Systems wird der Datenmigration nicht adäquat Rechnung getragen. Man stützt sich auf Tools ab und hofft auf die Lösung aller Probleme. Die Chancen eines datengetriebenen Ansatzes sind nicht bekannt oder werden nicht erkannt. Eine erfolgreiche Datenmigration ist in der überwiegenden Zahl der Fälle datengetrieben, der Einsatz von Tools optimiert die Bereitstellung der Daten auf der Zielumgebung.**

Bei der Überführung von Daten in eine neue Zielumgebung stehen die Entwickler meistens vor der Frage, „Wie ist die Ausprägung der im aktiven System vorhandenen Daten zu verstehen?“ Können die Daten, so wie sie vorhanden sind, also 1:1 übernommen werden oder sind hinter den Daten Prozesse vorhanden, die die Daten für den Benutzer umformen oder mundgerecht präsentieren, sei dies auf dem Bildschirm oder in Listen. Des weiteren ist die Herkunft oder Bereitstellung dieser Daten nicht immer bekannt. Insbesondere können diese aus verschiedenen Verarbeitungssystemen, mit unterschiedlichen Zeichensätzen (Code-Page), in das bestehende System übertragen worden sein oder der Benutzer machte sich neue Technologien zu nutze und hat die Daten mittels „cut and paste“ aus der Oberfläche seines Bildschirms in die Eingabefelder übertragen. Dies führt dazu, dass auf einmal in zu migrierenden Daten SteuerCodes wie „End of Line“ usw. in ganz normalen Textfeldern vorhanden sind. Solche Vorkommnisse sind ganz selten in den bestehenden Datenbeschreibungsdokumenten zu finden, wenn überhaupt solche vorhanden sind. Diese kleine Auswahl von Unstimmigkeiten zeigt auf, dass eine erfolgreiche Datenüberführung immer eine ganz intensive Auseinandersetzung mit den bestehenden Daten voraussetzt, zumal die Ersteller und Controller dieser Daten meistens nicht mehr im Betrieb verfügbar sind. Weitere Unannehmlichkeiten kommen

auf die Entwickler zu, wenn in der Vergangenheit aus Gründen der Kosten, des Aufwandes oder des Zeitfaktors, bei System- und Datenerweiterungen auf die Nachkonversion der bestehenden Daten verzichtet wurde. Unverträglichkeiten im bestehenden System wurden in diesen Fällen meistens über direkte „Datenmanipulationsprozesse“ korrigiert, zumal Auswertungstools wie SQL usw. nicht im Einsatz waren. Diese Ausgangslage zeigt auf, dass der Ansatz einer toolgetriebenen Datenmigration nicht zum gewünschten Ziel von vollständigen und sauberen Daten im neuen System führen wird. Der Ansatz muss im Kern, also bei den bestehenden Daten beginnen, was uns zu einer datengetriebenen Migration führt. Dieses Vorgehen ist als große Chance für die Unternehmung zu betrachten, das Wissen der Mitarbeiter um die Daten wird erhöht, und damit die Kompetenz für die Fehlerabklärungen beim Testen des neuen Systems sichergestellt.

## Vier Kernaufgaben einer datengetriebenen Migration:

### 1. Auseinandersetzung mit den vorhandenen Daten und Datenmengen

Datendokumentation konsultieren und wo sie fehlt, soweit bekannt, erstellen.

Bestehende, technische Datenstrukturen, die in Programmen verwendet werden, (COBOL-Copies,



PLI-%includes, DB-Definitionen/-Strukturen usw.) sammeln, analysieren, nachdokumentieren.

Wo notwendig, Prozesse analysieren und auf „Datenmanipulation“ überprüfen.

**2. Datenanalyse durchführen**

Maschinelles Abgleich der vorhandenen Daten mit der Beschreibung der Daten unter Benutzung von geeigneten oder zu erstellenden ABC-Analysen, unter Berücksichtigung der Zusammenführung von Systemen mit verschiedener Sprachherkunft.

Analysieren der Zeichensätze (Code Page), der Datenabweichung und deren Ursachenklärung.

Maschinelles Festhalten der Abweichungen und wo notwendig, manuelles Nachdokumentieren der Datenbeschreibungen.

**3. Datenveredelung analysieren, festlegen, ausführen**

Auf Grund der in der Datenanalyse festgestellten Unstimmigkeiten und deren Ursache, festlegen und dokumentieren von Datenveredelungsregeln, das heisst,

unter welchen Bedingungen wird ein bestehendes Datenelement, Datenfeld oder Datenattribut in eine für die Datenmigration oder das Mapping normierte Form gebracht, zum Beispiel: Umsetzung von EBCDICCode auf WE-ISO-8895-1.

Erstellen der Datenveredelungsmodule gemäß vorgegebenen, dokumentierten Regeln.

Erstellen des neuen, normierten Datensatzes, auf dem neben den eigentlich zu migrierenden Daten auch die Veredelungsregeln festgehalten sind und der nun Ausgangspunkt für das Mapping (Beschreibung der Datenmigrationsregeln) bildet.

**4. Durchführen der Datenmigration gemäß Vorgaben aus dem Mapping und ab der normierten Datenform aus der Datenveredelung**

Wo sinnvoll, laden der normierten Datenstrukturen aus der Datenveredelung in eine vorbereitete relationale Datenbank. Dies, um die eigentliche Datenmigrationstests effizient ausführen zu können. Da die Datenstrukturen normiert und die Zeichensätze standardisiert sind, kann das Laden unter Verwendung von geeigneten Tools ausgeführt werden.

Trennen der Daten in operative Daten und Daten des Rechnungswesens. Die operativen Daten folgen

speziellen Migrationsregeln, für die Überführung der Daten in das Rechnungswesen, zum Beispiel SAP, können entsprechende Anbieter-Tools gezielt eingesetzt werden.

Als geeignete Infrastruktur für die maschinelle Datenanalyse und die maschinelle Datenveredelung hat sich der Einsatz der Produkte von Micro Focus: „Micro Focus Net Express“ und Microsoft: „Windows XP“ erwiesen. Können doch auf diesen Plattformen zu analysierende und zu veredelnde Daten in sehr großen Datenmengen mehrere Zyklen auf dem PC durchlaufen ohne den oder die zentralen Rechner zu belasten. Dies erlaubt dem Entwickler saubere Analysen auszuführen und Testfälle zu simulieren. Durch einfache Transfers werden die so entwickelten Programme und Module

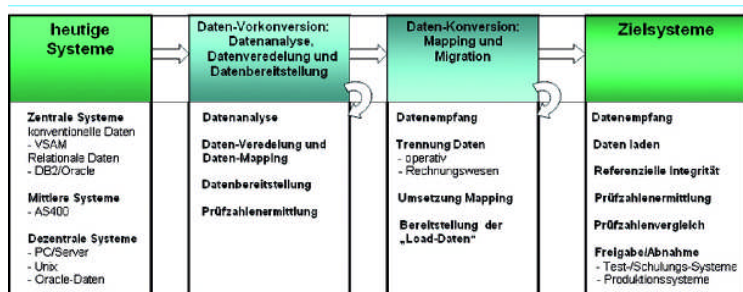
nach ausgereiften Tests auf den vorgesehenen Rechner transferiert, kompiliert und in das Save-Verfahren eingebunden. Dies bringt dem Betrieb den Vorteil, eine gezielte und synchronisierte Archivierung der bereitgestellten Daten

durchzuführen,

erleichtert das Wiederanlaufverfahren und stellt das Abnahmeverfahren der laufenden Migrationstests sicher. Durch die intensive Beschäftigung mit den bisherigen Daten steigt zudem das fachliche Wissen der eingesetzten Personen. Dieses kann wiederum zur Kontrolle der neuen Prozesse, zur Fehlerabklärung bei Unstimmigkeiten und für die Wartung und zukünftige Erweiterung des neuen Systems genutzt und zielorientiert eingesetzt werden.

Die vier Kernaufgaben der datengetriebenen Migration sind einfach in eine Verantwortlichkeitsmatrix und in eine Migrationsabwicklung zu integrieren und sichern dadurch ein effizientes und gezieltes Einsetzen von Tools bei der Migration sowie das Wissen der Datenverantwortlichen, erhöhen die Datenqualität und damit die sichere Einführung des neuen Systems.

Toolgetriebene Migrationen tun sich schwer mit der Realität der Daten, sie setzen zudem in den meisten Fällen voraus, dass alle Datenausprägungen definiert und auch so gespeichert sind. Die Unsicherheit über die Qualität der Daten bleibt weiterhin bestehen. ♦



Jürg Mathys, COMPORSYS Engineering AG, erreichbar unter [jmathys@comporsys.ch](mailto:jmathys@comporsys.ch)

