

COMPUTERWOCHE

NACHRICHTEN ♦ ANALYSEN ♦ TRENDS



Suchmaschine räumt SAP-Daten auf

*Mit einer Search Engine und Zusatzmodulen pflegt der TÜV Süd
1,1 Millionen Stammdatensätze in R/3 4.7.*



Eine Suchtechnik verhalf dem TÜV Süd zu Stammdaten ohne Dubletten.

Ein SAP-System einzuführen genügt noch lange nicht: Wenn ein Unternehmen Marketing- und Controlling-Instrumente effektiv nutzen möchte, müssen die Daten stimmen. So auch beim Dienstleister TÜV Süd. Nachdem die Datenstrukturen angepasst und das SAP-System entsprechend erweitert wurden, ge-

stalteten sich wichtige Auswertungen deutlich aussagekräftiger. Dabei hilft eine spezielle Suchtechnik, die den sekundenschnellen Zugriff auf den gesamten Adressbestand von etwa 1,1 Millionen Datensätzen erlaubt.

Doppelte Eingaben und falsche Schreibweisen

Das Problem von TÜV Süd dürf-

TÜV Süd in Zahlen

TÜV Süd ist ein technisches Dienstleistungsunternehmen mit den Geschäftsfeldern Industrie, Mobilität und Mensch. Das Leistungsspektrum der TÜV Süd Industrie Service GmbH umfasst Beratung, Prüfung, Tests und Gutachten in der Industrie. Die Spezialistenteams unterstützen Unternehmen bei Fragen zum sicheren und wirtschaftlichen Anlagenbetrieb sowie zum Arbeits- und Umweltschutz. Rund 2300 Mitarbeiter erwirtschafteten 2005 einen Umsatz von 260 Millionen Euro.

te vielen größeren Unternehmen bekannt sein: Durch Umorganisationen war ein gewaltiger Berg von Daten entstanden, der aus verschiedenen SAP-Systemen und anderen

Enterprise-Resource-Planning-(ERP-)Datenquellen stammte. Es war weder sicher, ob Datensätze mehrmals im System waren, noch, ob sie richtig geschrieben waren oder welche Relevanz sie hatten. Thomas Kainz, kaufmännischer Leiter von TÜV Süd, erinnert sich: „Wir konnten der Geschäftsleitung keine Daten liefern, die eine Zuordnung von Umsätzen zu bestimmten Kundengruppen zuließen.“ Außerdem konnten die Mitarbeiter keine Klammern um die Datensätze setzen, um vollständige Auswertungen vorzunehmen. Klammern sind im Fachjargon Beziehungen unter den Daten, zum Beispiel aus allen zu einem Konzern gehörenden Niederlassungen.

Die Daten des Prüfspezialisten entsprachen den Anforderungen der Vergangenheit, sollten aber dem aktuellen Bedarf eines modernen technischen Dienstleisters angepasst werden. Benötigt wurden auf den einzelnen Kunden oder auf Regionen bezogene Umsatzdaten. „Außerdem machten uns die Dubletten zu schaffen, die überall in den Stammdatensätzen auftauchten“, sagt Kainz. Erschwerend kam hinzu, dass durch die langen Prüfzyklen der Industrieanlagen ein Stammdatensatz, der die Adresse, den Ansprechpartner, Rechnungsadresse und die Umsätze enthält, oft nur einmal in fünf Jahren bearbeitet werden muss.

Erste Schritte mit einer Dublettenprüfung

Mehrere Versuche, den Datenbestand zu bereinigen, scheiterten an der Komplexität der Daten. Beispielsweise war es wichtig, Prüfobjekte, zu korrelieren wie etwa Aufzüge oder Gastanks, die sich oft an einem anderen



„Wir können jetzt unsere Umsätze genau nach Kundenunternehmen, Prüfobjekten oder Regionen aufschlüsseln.“

Thomas Kainz, kaufmännischer Leiter von TÜV Süd Industrie Service

Standort als dem Stammsitz des Betreibers befinden. Darüber hinaus waren viele Datensätze redundant vorhanden und blähten den Bestand unnötig auf. Kunden wie beispielsweise Großkonzerne bestehen aus vielen verschiedenen Betrieben und Niederlassungen, die teilweise die gleichen Daten melden. Zusätzlich erschwerten Franchise-Unternehmen die Sortierung, da sie zwar autark von einem Pächter betrieben werden, aber doch zu einem Verbund gehören.

Auf der Suche nach einem Werkzeug für die Bereinigung der Daten stieß Kainz auf die Firma ISO Software Systeme aus Nürnberg, die mit der Produkt-

reihe „Marble“ über die notwendigen Tools verfügt. Das Unternehmen integrierte seinerseits die Software „Fact-Finder“ von Omikron. Sie erlaubt eine unscharfe Suche innerhalb eines Datenbestandes. Dabei lassen sich ähnliche Begriffe besser finden als mit der Standard-Daten-

banksuche, die nicht fehlertolerant ist.

Einer der ersten Schritte war die externe Bereinigung der Adressdaten durch Omikron. Das Unternehmen erhielt dafür den kompletten Datenbestand. „Dies war die Feuerprobe, um zu erfahren, ob die Dublettenbereinigung funktioniert“, schildert Kainz. Zusätzlich zur Dublettenprüfung berichtete die Software die Postadressen, wobei falsche Postleitzahlen und vertippte Orts- sowie Straßennamen korrigiert und synchronisiert wurden. Die bearbeiteten Datensätze wurden in das SAP-System zurückgespielt. Seither stehen dem Sachbearbeiter in der Zentrale und dem Key-Account-Mitarbeiter vor Ort Klassifizierungsmerkmale zur Verfügung, die die alltägliche Arbeit erleichtern. So können Umsätze auf Objekte, Branche oder Konzernschlüssel bezogen werden.

Das Ergebnis der Adressbereinigung war beeindruckend: Von 723 406 Datensätzen wurden in über 130 000 Fällen die Adressen postalisch korrigiert. In knapp 20 000 Einträgen fehlte die Hausnummer. Weitere 3000 Datensätze waren entweder unvollständig oder fehlerhaft ausgefüllt. Damit waren gut ein Fünftel aller Daten nicht korrekt. Die Dublettenprüfung ergab, dass 83 000 Adressen doppelt vorhanden waren – also fast neun Prozent. In 640 000 Sätzen waren

SAP-Integration

Der TÜV-SÜD hat SAP R/3 Enterprise 4.7 im Einsatz und diese Lösung mit der Suchtechnik „**Fact-Finder**“ über Standardschnittstellen verbunden. Die Suchmaschine bedient sich eines indizierten Arbeitsspeichers, in dem die ERP-Daten abgelegt sind. Zugriffen wird auf die Daten der SAP-Software per **Remote Function Call (RFC)**. Der Fact-Finder ist in das für den Anwender gewohnte **SAP-Frontend** integriert. Über produkt-eigene Dynpros werden die Suchanfragen gestartet und ein Arbeitsvorrat gebildet.

Eine Besonderheit stellt der „Konzernmonitor“ dar: Mit ihm ist es möglich, **Kundenhierarchien** sowohl im Vertriebsinformationssystem (SAP-VIS) als auch im Controlling (CO-PA) zu strukturieren, und zwar in einem einzigen Arbeitsschritt. Normalerweise müssen Nutzer hierzu unterschiedliche **Transaktionen** bemühen sowie die modulspezifische Strukturierung der Daten berücksichtigen. Auf diese Weise und durch integrierte Funktionen wie „Drag and Drop“ sparen sich die SAP-Nutzer viel Detailarbeit. Den Konzernmonitor und die komplette Lösung implementierte ISO Software Systeme.

Zg	Objektnummer	Name	Name 2	Straße	PLZ	Ort	GF	Bez.	Konzern	Bezeichnung	HierZ
	0000300027	TÜV Süddeuts...		Westendstr.	80686	München			TÜV	TÜV Süd Konzern	
	0000300066	A+S s.r.o.		Malinovskyho 21	671 72	Miroslav					
	0000301652	TÜV Automot...		Westenstr.	80386	München			TÜV	TÜV Süd Konzern	
	0000301696	TÜV Manage...	TÜV SÜ...	Ridlerstr.	80339	München			TÜV	TÜV Süd Konzern	
	0000301703	TÜV Automati...	TÜV SÜ...	Westendstr.	80686	München			TÜV	TÜV Süd Konzern	
	0000301704	TÜV Akademi...		Westendstr.	80686	München			TÜV	TÜV Süd Konzern	

Debitor	Knoten	uooc	Name	Name 2	Straße	PLZ	Ort	GF	Bez.	Konzern
▼ TÜV										
▼ TÜV SÜD										
▼ SÜD	Süd									
0000301682			TÜV Rail Gmb...		Ridlerstr.	80339	München			TÜV
0000301712			TÜV Bayern G...		Westendstr.	80686	München			TÜV
▼ NORD	Nord									
0000301713			TÜV Immobili...	TÜV S...	Westendstr.	80686	München			TÜV
▼ WEST	West									
0000301618			TÜV Product ...		Ridlerstr.	80339	München			TÜV
0000301651			TÜV Verkehr ...		Ridlerstr.	80339	München	01	GmbH	TÜV
0000301604			TÜV Automoti...	TÜV S...	Westenstr.	80686	München	01	GmbH	TÜV
▼ TÜV GB1	Bed. Gebie...									

Konzernmonitor verbessert die laufende Datenpflege und Strukturierung im SAP-System.

dagegen keine Dubletten vorhanden. Die Dublettenliste reichte von phonetisch identischen Eingaben wie beispielsweise Mayr oder Meier über Umzugsdubletten bis hin zu Abweichungen im Firmennamen wie etwa Garten-Meier und Meier Gartenbau.

Aber die Bestandsbereinigung war für Kainz nur der erste Schritt. Zusammen mit ISO Software Systeme wurde der „Konzernmonitor“ projektiert, der die laufende Datenpflege und Strukturierung im SAP-System verbessert. Er nimmt die Dublettenprüfung im Dialog und im Bestand (Batch) vor, bietet eine fehlertolerante Kundensuche (Debitor) und unterstützt die Adressvalidierung. Insgesamt 1,1 Millionen Stammdatensätze werden ständig durchforstet. 700 000 Einträge sind umsatztragend und haben eine ausgeprägte Buchungskreissicht, damit zum Beispiel der Geschäftsprozess der Rechnungslegung (Faktura) abgebildet werden kann. 180 000 Daten werden jedes Mal im Rahmen der jeweiligen Prüfzyklen gepflegt.

Sachbearbeiter können Datensätze leichter anlegen

Die Pflege der Daten hat sich gegenüber dem regulären SAP-

System wesentlich verbessert. Der Sachbearbeiter kann jetzt weitaus einfacher Daten anlegen und strukturieren. Der Aufbau einer Kundenstruktur erfolgt analog dem Erstellen einer Baumstruktur im Microsoft-Explorer. Drag and Drop wird dabei im Konzernmonitor ebenfalls unterstützt. Die Stammdaten erhielten einen Konzernschlüssel, um die einzelnen Angaben besser zuordnen zu können. Die Strukturen für die Ergebnisrechnung auf Geschäftsleitungsebene (SAP-Modul CO-PA) und für das Vertriebsinformationssystem (SAP-Modul VIS) werden automatisch synchron aktualisiert.

Außerdem durchforsteten die Mitarbeiter die Stammdaten über eine intelligente Suchtechnik, die auf Fact-Finder basiert. Das Auffinden von Adressen wurde durch diese unscharfe Suche wesentlich erleichtert. Anstatt wie früher Wildcards, etwa „Me?er“ oder „Garten*“, eingeben zu müssen, können die Mitarbeiter nun eine beliebige Schreibweise verwenden und erhalten als Ergebnis alle passenden Datensätze, die ähnlichsten zuerst. „Diese SAP-Anpassung erforderte einige Schleifen mit dem Systemintegrator ISO Software Systeme, war aber durch die parallele Strukturierung der

Daten absolut im Rahmen unseres Projektes“, so Kainz.

350 Anwender arbeiten mit dem Such-Server

Insgesamt rechnet der kaufmännische Leiter 150 000 Euro für die Umsetzung. Der interne Aufwand von vier bis fünf Fachleuten schlägt zusätzlich mit etwa 70 000 Euro zu Buche. Rund 350 lizenzierte Arbeitsplätze sind an den Such-Server angeschlossen. Wie stark die Nutzung und die Akzeptanz sind, zeigte ein Stromausfall, nach dem erst langsam alle Server wieder hochgefahren wurden. „Sofort riefen die Nutzer beim Support an und fragten nach der Verfügbarkeit der Suchtechnik“, erinnert sich Kainz. Das SAP-System wird vom eigenen IT-Dienstleister TÜV Süd Informatik und Consulting Services GmbH betrieben. Für die fehlertolerante Suche mit 1,1 Millionen Stammdaten werden etwa 300 MB Arbeitsspeicher benötigt. Als Zentralrechner fungiert in der Regel ein SAP-Applikations-Server oder, wie beim TÜV Süd, ein eigener Dell-Server.

Gezielte Mitarbeiterschulung in zwei Schritten

Die ersten Erfahrungen zeigten, dass es für die Mitarbeiter wich-

tig ist, die Suche im Konzernmonitor mit bestimmten Kniffen zu verfeinern, um bessere Ergebnisse erzielen zu können. Die Datenpfleger bekamen dafür ein Training. In einer zweiten Schulung wurde die Bedienung des Konzernmonitors geübt. Mittlerweile sind die Verbesserungen in den Alltag eingeflossen. Dass die Suche ständig 1,1 Millionen Datensätze durchforstet, wird den Mitarbeitern angesichts der schnellen Antwortzeiten nicht bewusst. Weniger gute Erfahrungen machte Kainz mit der Organisation der Stammda-

teneingabe: Die Fehlerquote von externen Hilfskräften war so hoch, dass Kainz auf die eigenen Mitarbeiter zurückkam. Die Sachbearbeiter aus dem Bereich Rechnungswesen lieferten die besten Ergebnisse.

„Wir sind sehr zufrieden mit dem Projekt“, bilanziert Kainz. Die Außendienstmitarbeiter können genaue Anfragen zu Kunden stellen. Außerdem operieren alle Nutzer mit den gleichen Daten. Beispielsweise lassen sich schwierige Strukturen wie etwa Franchise-Unternehmen einfach darstellen.

In puncto weiterer Ausbau ist geplant, die Verbesserungen auch in anderen SAP-Anwendungen einzuführen. Um die Datensätze noch schneller zu bearbeiten, werden künftig auch Löschvorkennungen oder Buchungskreisdaten in die Suche einbezogen. Kainz resümiert: „Der korrekte Datenbestand ist eine optimale und notwendige Ausgangsbasis für unseren Vertrieb.“ (fn) ◆

***HANS-JÖRG SCHILDER** ist freier Redakteur in Wessobrunn.