

Search Engine Cleans Up SAP Data

Using a search engine and add-on modules, TÜV Süd maintains 1.1 million master data records in SAP R/3 4.7.



Search technology assisted TÜV Süd in the quest for duplicate-free master data.

If a company wants to use marketing and controlling mechanisms effectively, the data must be right – and implementing an SAP system is simply not enough. Technical service company TÜV Süd, headquartered in Munich, made this discovery too. After the data structures had been adjusted and the SAP system enhanced, the information delivered by key reports became much more reliable. Special search technology helps here by enabling access to the entire address list of around 1.1 million data records within seconds.

Duplicate entries and incorrect spelling

The problems encountered by TÜV Süd will, no doubt, be familiar to many other large companies too. As a result of restructuring, a mountain of data was generated from different SAP systems and other enterprise resource planning (ERP) data sources. No-one was sure whether the data records were duplicated in the system, whether the spelling was correct, or how relevant they were. Thomas Kainz, business manager at TÜV Süd, explains, "We couldn't supply the board with data that enabled them to assign sales to certain customer groups."

TÜV Süd in figures

TÜV Süd is a technical service provider active in the business segments industry, mobility, and people. Its range of services embraces consultancy, inspections, tests and surveys for industry. TÜV Süd's teams of specialists help companies run their installations securely and economically, and provide services in the areas of occupational safety and environmental protection. In 2005, around 2,300 employees generated sales of EUR 260 million.

What's more, employees weren't able to compound data records to run complete reports (compounding describes relationships in the data, such as all subsidiaries that belong to a group).

TÜV Süd's data met past requirements but had to be adapted to meet the current needs of a modern technical service provider. What the company needed was sales data related to individual customers or regions. "In addition, duplicate records were a headache – they were everywhere in the master data records," says Kainz.



“We can break down our sales figures exactly according to company, inspection object or region.”

Thomas Kainz, business manager at TÜV Süd Industrie Service

Another difficulty was that, due to the long inspection cycles of industrial installations, a master data record containing the address, contact person, invoice address, and sales data often had to be processed just once every five years.

Initial steps with a duplicate check

Several attempts to cleanse the dataset failed because the data was too complex. For example, it was important to correlate inspection objects such as elevators or gas tanks, which are often not located in the same place as the operator’s headquarters. In addition, many data records were redundant and inflated the amount of data unnecessarily. Customers such as major conglomerates comprise many different companies and subsidiaries, some of whom report the same data. Furthermore, franchise companies made sorting the data more difficult, because they are operated independently by a franchisee, but still belong to a group.

In the quest for a tool to cleanse the data, Kainz came across the Nuremberg-based company ISO

Software Systeme, whose Marble product range offered the perfect solution. In turn, ISO Software Systeme integrates the FACT-Finder software from Omikron, based in the German town of Pforzheim. FACT-Finder enables fuzzy searches to be made within a dataset. This means that similar terms can be found more easily than with the standard database search, which is not error-tolerant.

One of the first steps involved the external cleansing of the address data by Omikron. The company was supplied with the complete dataset for this purpose. “This was the acid test to find out whether duplicate cleansing worked,” says Kainz. In addition to checking for duplicates, the software amended postal addresses by correcting and synchronizing incorrect post codes and misspelled town and street names. The processed data records were then imported back into the SAP system.

Employees at head office and key account employees on site can now use classifying attributes that make their day-to-day work easier and can generate sales figures based on individual items, industry sectors, or corporate group keys.

The address cleansing results were impressive. From a total of 723,406 data records, 130,000

SAP integration

TÜV Süd deploys SAP R/3 Enterprise 4.7 and has integrated it with FACT-Finder search technology using standard interfaces. The search engine uses an index-based working memory in which the ERP data is stored. The SAP software data is accessed by Remote Function Call (RFC). FACT-Finder is integrated into the SAP front end, which users are already familiar with. Search requests are started using FACT-Finder’s own screens and a worklist is created.

The Group Monitor is a special feature: It enables users to structure customer hierarchies in the Sales Information System (SAP SIS) and in Controlling (CO-PA) in a single step. Users would normally have to access different transactions to do this, and would also have to take account of how the data is structured in each component. As a result – and with integrated functions like drag-and-drop – the SAP users save themselves a lot of time and effort. ISO Software Systeme implemented the complete solution and the Group Monitor.

Zg	Objektnummer	Name	Name 2	Straße	PLZ	Ort	GF	Bez.	Konzern	Bezeichnung	Herz
	0000300027	TUV Süddeuts		Westendstr	80666	München			TUV	TUV Süd Konzern	
	0000300066	A+S s.r.o		Milivojeva 21	671 72	Miroslav					
	0000301652	TUV Automab		Westendstr	80386	München			TUV	TUV Süd Konzern	
	0000301696	TUV Manag	TUV SU...	Ridderstr	80339	München			TUV	TUV Süd Konzern	
	0000301703	TUV Automab	TUV SU...	Westendstr	80666	München			TUV	TUV Süd Konzern	
	0000301704	TUV Akademi		Westendstr	80666	München			TUV	TUV Süd Konzern	

Debitor	Knoten	jurrt.	Name	Name 2	Straße	PLZ	Ort	GF	Bez.	Konze
			TUV							
			TUV SÜD							
			SÜD	Süd						
			0000301682	TUV Rail Gmb...	Ridderstr.	80339	München			TUV
			0000301712	TUV Bayem G...	Westendstr.	80666	München			TUV
			NORD	Nord						
			0000301713	TUV immobile...	TUV S...	Westendstr.	80666	München		TUV
			WEST	West						
			0000301618	TUV Product ...	Ridderstr.	80339	München			TUV
			0000301651	TUV Verkehr ...	Ridderstr.	80339	München	01	GmbH	TUV
			0000301694	TUV Automot... TUV S...	Westendstr.	80666	München	01	GmbH	TUV
			TUV GB1	Bed. Gebie...						

The Group Monitor improves ongoing data maintenance and structuring in the SAP system.

postal addresses were corrected. In almost 20,000 entries, the house number was missing. A further 3,000 data records were filled out incorrectly or were incomplete. This means that one fifth of all the data was incorrect. The duplicate check revealed that 83,000 addresses had been entered twice in the system, which amounts to 9 per cent. Around 640,000 records, on the other hand, were not duplicated. The list of duplicates included phonetically identical entries such as Mayr and Maier, which are pronounced the same in German, duplicates as a result of a change of address, and variations in company names, such as “Meier Gardens” and “Meier’s Market Gardens”.

But cleansing the dataset was just the first step for Kainz. Together with ISO Software Systeme, TÜV Süd set up the Group Monitor, which improves ongoing data maintenance and structuring in the SAP system. It checks for duplicates online and in the background, provides an error-tolerant search function for customers, and supports address validation. A total of 1.1 million master data records are now scanned on a continuous basis. 700,000 of the entries are sales-related and have a detailed company-code view enabling the business process

for accounting to be mapped, for example. 180,000 records are maintained in each inspection cycle.

Easier data record creation

Data maintenance has improved greatly compared with the regular SAP system. Employees can now create and structure data much more easily. They create a customer structure in the same way as they would a tree structure in the Microsoft Explorer. Drag-and-drop is also supported in the Group Monitor. The master data was allocated a corporate group key so that the individual details could be assigned more effectively. In addition, the structures for Profitability Analysis at board level (SAP component CO-PA) and for the Sales Information System (SAP component SIS) are automatically updated at the same time.

Furthermore, employees search the master data using intelligent search technology based on FACT-Finder. With this fuzzy search, it is now much easier to find addresses. Instead of having to type in wildcards, such as “Me?er” or “Garden*”, employees can spell the word as they choose and the system delivers all suitable data records – with the most similar ones first. “Adapting the SAP system took a

350 users work with the search function

According to Kainz, the implementation cost some EUR 150,000. The internal cost of four to five experts adds another EUR 70,000 to the bill. Around 350 licensed users are connected to the search server. A power cut in which all the servers had to be restarted showed just how extensively the solution is used. "The users phoned the support hotline immediately and asked when the search functionality would be available again," says Kainz. The SAP system is operated by TÜV Süd's own IT service provider TÜV Süd Informatik und Consulting Services GmbH. For an error-tolerant search with 1.1 million master data records, around 300 MB working memory is required. The application usually runs on an SAP application server or – in the case of TÜV Süd – a dedicated Dell server.

Two-step targeted training

Initial experience showed that, to obtain better results, it is important for employees to refine the search in the Group Monitor using certain tricks. Those responsible for maintaining data were given specific training for this purpose. In a second training course, employees practiced using the Group Monitor. In the meantime, the improvements have been incorporated into everyday working life. Due to the rapid response times, employees are not aware that the search function constantly scans 1.1 million data records. However, Kainz was not as satisfied with how master data entry is handled. The number of errors made by external temporary workers was so high that Kainz had to entrust TÜV Süd's regular staff with the task. Employees from the accounting department produced the best results.

"We are very happy with the project," Kainz sums up. Field employees can contact customers with exact inquiries. What's more, all users work with the same data. In addition, complicated structures such as franchises can be mapped easily.

As far as the future is concerned, TÜV Süd plans to implement the improvements in other SAP applications too. To process the data records even more quickly, deletion flags or company code data will also be included in the search function. Kainz concludes, "A correct dataset is an optimum and necessary starting point for our sales division." (fn)

*HANS-JÖRG SCHILDER is a freelance editor in Wessobrunn, Germany.