

VAN DER LANS

# Een nieuwe categorie BI-producten?



Jaren geleden waren gebruikers blij als ze een simpel rapport op hun monitor of PC konden opvragen. Maar al snel bleek dat niet voldoende. Gebruikers wilden over OLAP-mogelijkheden beschikken om dynamisch gegevens te bestuderen, ze hadden hun zinnen gezet op het maken van complexe statistische modellen om de toekomst te voorspellen en ook moesten er dashboards komen die inzicht zouden geven in de bedrijfsprocessen. Het effect is dat er nu veel categorieën BI-producten bestaan waarmee gebruikers de gegevens kunnen bestuderen.

Het lijkt alsof we er een nieuwe categorie bij krijgen, de zogenaamde *data discovery* producten; ook wel *data exploration* producten genoemd. Bij veel van de andere categorieën zijn de gebruikers, voor wat betreft het stellen van vragen, beperkt door de structuren in de data en door de gedefinieerde relaties tussen tabellen. Bijvoorbeeld, als de technici zich niet gerealiseerd hebben dat de gebruikers op een bepaalde manier van de ene naar de andere tabel willen navigeren, dan zal het ook nagenoeg onmogelijk zijn voor de gebruikers om dit pad te volgen.

Data discovery producten staan gebruikers volledig vrij door de gegevens heen te navigeren zonder enkele beperking. Relaties kunnen gelegd worden waar de technici niet aan gedacht hebben. Dit soort functionaliteit is uitermate handig voor die problemen waarvan het initieel niet duidelijk is waar het antwoord gezocht moet worden. Als we bijvoorbeeld willen weten hoe de omzet er de laatste maanden uitzag, dan weten we welke tabellen we moeten benaderen en hoe. Maar wat als we weten dat er een probleem is, maar we geen idee hebben van wat de oorzaak is? Stel bijvoorbeeld dat de uitval in een fabriek aan het toenemen is en we willen weten waarom. Waar ligt dan het probleem? Afnemende kwaliteit van de aangeleverde grondstoffen? Bepaalde machines die vaker in de fout gaan? Minder ervaring bij de werknemers die de machines bedienen? Of wat als de verkoop van bepaalde producten in een regio tegenvalt? Is de toelevering van nieuwe producten te traag? Is er misschien een staking geweest? Is het weer zo slecht dat potentiële klanten amper hun huis uitkomen?

De intentie van data discovery producten is dat gebruikers vrij vragen mogen stellen. Er zijn nagenoeg geen beperkingen wat betreft het stellen van vragen. Samen vormen deze producten echter geen homogene verzameling producten, zoals alle SQL databaseservers dat wel doen. Ze proberen het allemaal op hun eigen wijze op te lossen.

Een interessant data discovery product is iCorrelate van illuminate. Dit product biedt de gebruiker een simpele grafische interface waarmee vrij vragen gesteld kunnen worden. Bijvoorbeeld, in iCorrelate kunnen we de vraag stellen: Zoek Rotterdam. Het antwoord zal bestaan uit alle records uit alle tabellen waar ergens het woord Rotterdam in voorkomt. Hierna kan gevraagd worden welke records dan te maken hebben met klant Jansen. Het antwoord kan klachten van deze klant bevatten, maar ook bestellingen en facturen. Het product gaat hierbij uit van een dialoog; gebruikers kunnen vragen stellen op antwoorden van voorgaande vragen. iCorrelate onthoudt alle antwoorden in een sessie.

Kortom, de gebruiker kan vrij door de database heen wandelen, zonder dat ook maar één relatie vooraf gedefinieerd is.

## Er zijn nagenoeg geen beperkingen wat betreft het stellen van vragen

Een geheel ander product is I.Know van InterSystems. Puur door middel van tekstanalyse kunnen gebruikers vrij tekst op een gestructureerde manier bestuderen en analyseren. Een derde die ik hier wil noemen is Exalead. Dit product zet search-technologie en natuurlijke taal in om gebruikers vrij te laten zoeken. Gebruikers kunnen vrij data in gestructureerde bronnen combineren met data in ongestructureerde bronnen.

Er zullen ongetwijfeld nog meer producten bestaan die tot deze categorie behoren en er zullen er nog diverse verschijnen, want de categorie begint 'hot' te worden. Zeker nu Gartner erover publiceert. Als u zich hier nog niet in verdiept hebt, kan ik het wel aanraden. Data discovery of data exploration is een waardevolle aanwinst voor de Business Intelligence wereld.

Opmerking: Verwar data discovery niet met de term knowledge discovery. Laatstgenoemde term wordt momenteel in de BI-wereld niet veel meer gebruikt, maar enkele jaren geleden wel. Het werd toen gebruikt als een alternatief voor data mining.

**Rick van der Lans** is zelfstandig IT-consultant.