

Integratie van ongestructureerde data wordt cruciaal

Trends en ontwikkelingen in Business Analytics

Mirjam Hulsebos

Business Analytics werd tot op heden vooral gebruikt om gestructureerde interne informatie te analyseren, zodat ontwikkelingen in het verleden konden worden verklaard. Maar het is vaak juist door het gebruiken van externe en ongestructureerde informatie dat kansen voor verbetering zichtbaar worden.

Niet voor niets hebben bedrijven dan ook steeds meer behoefte aan voorspellende analyses in plaats van beschrijvende. SAS ziet vier belangrijke ontwikkelingen in de manier waarop data beschikbaar komen en worden gebruikt. DB/M woonde onlangs het jaarlijkse klantenevenement van SAS bij. Deze leverancier van Business Analytics investeert momenteel flink in allerlei vormen van analyse van ongestructureerde data, van gesprekken in het contactcenter tot en met het in kaart brengen van sociale netwerken van consumenten.

Snellere analyse met meer data

De eerste ontwikkeling is dat organisaties steeds sneller informatie beschikbaar moeten hebben. Tegelijkertijd hebben ze te maken met een enorme hoeveelheid aan beschikbare data. High Performance Analytics maakt het mogelijk om 'big data' in heel korte tijd te analyseren. In het verleden gebruikte men bij grote databestanden vaak slechts een subset van de data, omdat anders de berekeningen te lang duurden. SAS heeft In-Database- en In-Memory Analytics-technologie ontwikkeld. Door het analyseproces in de database zelf te brengen (in-database), en met minimalisatie van I/O-processen (in-memory), kunnen met grote snelheid, grote hoeveelheden data geanalyseerd worden. Dus niet op een gerepliceerde subset, maar direct op de gehele dataset. Een stresstest bij banken bijvoorbeeld wordt met deze technologie van weken verkort naar een dag. SAS voerde onlangs met zijn High Performance Risk-software, die gebruik maakt van In-Memory Analytics, 8,8 miljard value at risk-calculaties uit bij een financiële instelling. Waar deze calculaties normaal op één server 18 uur in beslag zouden nemen, werden dezelfde berekeningen nu binnen drie minuten gerealiseerd. Ook voor processen in andere sectoren heeft SAS deze technologie in specifieke oplossingen toegepast. Daarnaast is SAS een samenwerkingsverband aangegaan met

zowel Teradata als EMC. Dit betekent dat SAS High Performance Analytics beschikbaar is op een Teradata architectuur en op de EMC Greenplum Data Computing Appliance. Deze combinatie maakt het mogelijk om complexe berekeningen parallel te draaien in een grid. De verschillende jobs worden verdeeld over verschillende virtuele servers en parallel uitgevoerd. Dat brengt de tijd die nodig is voor berekeningen nog verder terug, van uren naar minuten of zelfs seconden.

Meer verschillende bronnen

De tweede trend is dat niet alleen de hoeveelheid data toeneemt, maar ook het aantal verschillende bronnen dat voor een organisatie belangrijk wordt. Werden vroeger vooral interne gestructureerde databronnen gebruikt voor analyses (bijvoorbeeld uit transactionele systemen), nu komen daar ook externe bronnen bij en deze zijn vooral ongestructureerd. Met name als het gaat om marketingtoepassingen is het aantal databronnen haast onuitputtelijk, denk aan de mening die consumenten op fora of vergelijkingssites spuien, berichten op Twitter en Facebook, gesprekken in het contactcenter en zoekgedrag op websites (social media Analytics). Maar denk bijvoorbeeld ook aan sensoren die automatisch allerlei gegevens verzamelen voor monitoringdoeleinden, zoals sensoren op verpakkingendozen of rolcontainers die in de gaten houden of de temperatuur tijdens transport niet te hoog is geweest; sensoren die monitoren wat het loopgedrag van koeien in een ligboxenstal is; of ongediertebestrijdingbedrijven die op afstand kunnen zien of muizen vallen al dan niet zijn dichtgeklapt. Ook intern neemt het aantal bronnen toe. Door fusies, overnames en innovaties voeren de meeste bedrijven vandaag de dag verschillende productlijnen en merken, vaak ondergebracht in meerdere dochterbedrijven die allemaal met hun eigen software werken (zie kader Vail Resorts). Deze ontwikkelingen maken het noodzakelijk om alle beschik-

bare data te integreren, te beheren, en op een bepaald kwaliteitsniveau te houden. Het belang van datamanagement wordt dus steeds groter. SAS heeft haar datamanagement platform dan ook ingericht om alle benodigde elementen te ondersteunen, van data-integratie en datakwaliteit tot masterdata management.

Integratie in processen

De derde trend is dat analyses niet in een losstaand proces tot stand komen. De effectiviteit van de analyse van data wordt namelijk hoger als deze systematisch geïntegreerd is in bedrijfsprocessen, denk aan de medewerker van het ongediertebestrijdingbedrijf die op zijn PDA doorkrijgt op welke locaties hij nieuwe muizenvallen moet plaatsen; of een boer die een SMS krijgt als één van de koeien de hele dag nog niets heeft gegeten en ook nog niet in de melkmachine is geweest. Door de integratie van Analytics in processen wordt het onder meer mogelijk om kwaliteitscontroles in het proces aan te brengen. Maar het kan ook gaan om analyses met een voorspellende waarde. Een voorbeeld daarvan is het optimaliseren van een schappenplan van een supermarkt. Als consumenten door een bacterie plotseling geen komkommers meer kopen, dan moet het versvak anders worden ingericht. Wat is dan de nieuwe optimale schapindeling waarbij de omzet wordt gemaximaliseerd? Waar dit soort analyses tot voor kort vaak wel op het hoofdkantoor kon worden gemaakt voor grote groepen supermarkten, is het nu mogelijk om per dag en per supermarkt de schapindeling te optimaliseren, rekening houdend met lokale voorkeuren. Dit soort informatie is bovendien niet alleen op de computer in het kantoorje toegankelijk, maar ook op mobiele devices. Dan kan de supermarkt manager voor het schap gaan staan en op zijn PDA, smartphone of tablet zien wat de snel lopers en langzaam lopers zijn en bij welke schapindeling de omzet wordt gemaximaliseerd. SAS komt later dit jaar met een eigen workflow engine, die kan worden geïntegreerd in BPM-tools. Daarmee wordt het mogelijk om Analytics volledig in bestaande processen te integreren.

Democratisering van informatie

De vierde trend is het toegankelijk maken van Analytics voor de business, ook wel het democratiseren van Analytics genoemd. Gebruikers van informatie kiezen zelf hoe en welke informatie ze tot zich willen nemen. "Power to the people", zo vat één van de sprekers op SAS Forum het samen. "Geef medewerkers op de werkvloer toegang tot Analytics zodat zij het terugkerende werk zelf kunnen doen. Bundel de analytische kennis in een Centre of Excellence, dat de handen vrijhoudt voor complexere vraagstukken en dat bovendien een brug kan slaan tussen business en IT." Om Analytics op de werkvloer toe te passen is het belangrijk dat de informatie 'fit for task' is, zoals het eerder genoemde supermarktvoorbeeld, waarbij de winkelmanager een schappenplan op zijn mobiele device ziet. Vermoei medewerkers niet met ellenlange rapportages, maar geef ze alleen die informatie die ze nodig hebben om op dat moment hun werk beter te doen. Denk bijvoorbeeld aan een account manager die vlak voor een klant-

In vijf stappen naar Data Management

Data Management, ofwel de processen en technologie om data in verschillende systemen op een eenduidige manier vast te leggen, te integreren en te beheren, staat sterk in de belangstelling. Geen wonder, want het aantal databronnen dat wordt gebruikt bij analyses groeit sterk. Hoe meer systemen en bronnen, hoe lastiger het is de integriteit te bewaken en hoe sterker de roep om compliancy. Het belang van consistente data is de afgelopen jaren immers alleen maar toegenomen. Vandaar ook de wens om de verschillende bronnen te integreren in één systeem, met voldoende maatregelen om de integriteit en volledigheid van de data te garanderen. Dat kan in vijf stappen:

1. Speel in elkaars zandbak. Overstijg de afdelingsgrenzen en kijk in andere business units of bij andere dochterondernemingen hoe zij omgaan met data. Breng de kennis, vaardigheden, cultuur en technologie van al die verschillende business units op hetzelfde niveau;
2. Schoon de data op. Garbage in is garbage out, dus zorg dat de data aan de bron schoon zijn, d.w.z.; accuraat, integer en compleet;
3. Democratiseer Analytics, laat iedereen ermee werken. Je kunt Analytics in bijna ieder proces gebruiken om de toekomst te voorspellen of om te signaleren dat iets niet naar wens verloopt. Laat de business zelf bedenken hoe hun proces kan worden verbeterd door meer en beter gebruik te maken van de beschikbare informatie;
4. Stel regels op en volg ze. Data Governance formaliseert het proces van het beheren van informatie dwars over alle afdelings- en business unit-grenzen heen. Het behoeft geen betoog dat het belangrijk is dat iedereen in het bedrijf daarbij dezelfde regels volgt;
5. Schaf de juiste tools aan. In de eerste vier stappen heeft u ervoor gezorgd dat er een cultuur ontstaat waarin medewerkers het belang van het hebben van informatie inzien en er ook voor zorgen dat die informatie compleet en integer is. De laatste stap tot slot is het analyseren van die data zodat de daaruit voortvloeiende informatie kan worden ingezet om betere beslissingen te nemen.

bezoek op zijn smartphone nog even snel de omzet in de afgelopen periode, de klachten die de klant heeft geuit en de hoogte van de openstaande facturen inziet. En dat alles niet via drie verschillende rapporten, maar gewoon op één scherm zodat het simpel en overzichtelijk is. SAS gaat mee in de ondersteuning van deze democratisering door Business Analytics nog toepasselijker voor iedereen te maken. De software is op elk mobiel apparaat te gebruiken en door Outlook-integratie worden analyses direct in de meest gebruikte applicatie, die voor e-mail, beschikbaar gesteld. Op die manier is Business Analytics niet langer een speeltje van de directie, maar wordt het een stuurinstrument voor alle medewerkers.

Vail Resorts verbetert klantervaring met Analytics

Hoe kun je de beleving van een skivakantie optimaliseren? Hoe zorg je ervoor dat een skigebied langer interessant blijft voor gasten van je hotel? En hoe kun je met een kleiner marketingbudget meer mensen bereiken? Vail Resorts vond het antwoord in Customer Intelligence.

Vail Resorts is een Amerikaanse keten die alles rondom skivakanties organiseert: van verblijf tot transport en van lessen tot het verhuur van materiaal. Zij behoren tot de pioniers van het concept experience marketing: al jarenlang focust het bedrijf zich op het bieden van een optimale klantervaring. Het beheren, verbeteren en exploiteren van de klantendatabase is daarbij cruciaal, vertelt Tim April, vice president Information Technology. "We zijn gegroeid door acquisities en overnames en hebben verschillende lines of business. Aan de achterkant hebben we dus vele verschillende CRM-, hotelmanagement- en andere systemen met klantinformatie. Bovendien werken we ook veel samen met partners, die natuurlijk hun eigen software hebben. Het is een uitdaging om al die data samen te brengen, en een nog grotere uitdaging om die informatie te gebruiken om klanten een naadloze klantervaring te bieden. Want het gaat er natuurlijk niet om zoveel mogelijk klantinformatie te verzamelen, het gaat erom informatie te gebruiken die een voorspellende waarde heeft voor klantgedrag in de toekomst."

Om te analyseren welke klanten het best met welk aanbod kunnen worden benaderd, gebruikt het bedrijf SAS Customer Intelligence. "Tot voor kort was onze marketing productgedreven aangevuld met intuïtie. Sinds we analytische software gebruiken om te identificeren bij

welke klanten de responskans het hoogst is, benaderen we klanten veel effectiever. Het effect van onze marketinginspanningen is daardoor nu een stuk hoger."

Sinds deze winter heeft Vail Resorts er een belangrijke bron van klantdata bij: informatie van het liftensysteem. Het liftbedrijf is een dochter van Vail Resorts en verzamelt al enkele jaren informatie over welk pasnummer welke liften neemt. Afgelopen winter koppelde Vail Resorts het pasnummer voor het eerst aan de klant waardoor het bedrijf precies kan zien wie welke routes skiet. Ze ontdekten welke gasten altijd alleen de voor hen vertrouwde pistes nemen en wie er open staat voor experimentele routes. Het bleek bijvoorbeeld dat de eerste groep sneller is uitgekeken op een skigebied en dus minder vaak terugkomt. Door dit soort analyses ontstonden tientallen ideeën om de omzet te verhogen.

Maar nog revolutionairder was het idee om deze informatie ook met klanten te delen, op zo'n manier dat dit bijdraagt aan hun ski-ervaring. Zo ontstond EpicMix, een website waarop gasten kunnen zien welke routes ze hebben geskied, hoeveel hoogtemeters ze hebben gemaakt en hoe lang ze hebben geskied. Ze kunnen ervoor kiezen deze informatie te delen met anderen, niet alleen via EpicMix maar ook via Facebook. Door één simpel vinkje verschijnt alle informatie real-time op je Facebookpagina. Ga je op skivakantie, dan kun je via EpicMix zien welke routes andere mensen skieën. "Dat stimuleert gasten om zelf nieuwe routes te verkennen. Zo blijft een skigebied langer interessant en neemt de kans toe dat mensen nog eens terugkomen", geeft April als voorbeeld.

Social media Analytics

Een terrein waarvan steeds meer organisaties inzien dat ze er niet aan voorbij kunnen gaan, is social media Analytics: wat wordt er over het bedrijf of merk gezegd op Twitter, Facebook of op recensiewebsites? Gecombineerd met predictive modeling levert dit een zeer krachtige bron van stuurinformatie op.

Keith Collins, CTO bij SAS, geeft een voorbeeld: "Er is een Amerikaans ziekenhuis dat met behulp van social media analysis en predictive modeling griepuitbraken voorspelt. Zij scannen Twitter, Facebook en andere social media sites op woorden als griep of verkoudheid en kunnen op basis van het aantal keren dat zo'n woord voorkomt voorspellen of en waar er een griepuitbraak komt." Een ander voorbeeld is een garagebedrijf dat aantekeningen van monteurs digitaal vastlegt en analyseert. Collins: "Voorheen registreerde dat bedrijf alleen welk werk de monteur had verricht om op basis daarvan de factuur te kunnen maken. Nu vragen ze de monteurs ook te noteren wat ze opvalt aan een auto, bijvoorbeeld een zwak punt waar ze op dat moment echter geen reparatie aan hebben verricht. Zo krijgen ze meer inzicht in de kwaliteit van auto's en kunnen ze beter voorspellen welke onderdelen het snelst kapot zullen gaan."

De onderliggende taxonomie kan iedere organisatie zelf aanpassen, afgestemd op het taalgebruik, de woordkeus en de leeftijd

van de doelgroep. Uiteraard is het mogelijk dit soort bronnen van ongestructureerde informatie te combineren met gestructureerde klantdata uit onder meer het CRM-systeem.

Conclusie

In vrijwel ieder proces kun je analytics gebruiken om de toekomst te voorspellen of om te signaleren dat iets niet naar wens verloopt. Waar Business Analytics tot voor kort eigenlijk alleen kon worden toegepast op gestructureerde interne databronnen, is het nu ook mogelijk om externe bronnen en vooral ongestructureerde bronnen mee te nemen in de analyse. Dat betekent natuurlijk wel dat de te analyseren hoeveelheid informatie drastisch toeneemt. Maar door nieuwe technologieën als In-Memory Analytics kunnen de meeste berekeningen toch nog altijd vrijwel real-time worden uitgevoerd. En dat is belangrijk, want organisaties willen snel op ontwikkelingen inspelen. Daarnaast moet Business Analytics steeds dieper in hun processen integreren: geen maandelijks duimendik rapport maar 'on demand' en 'fit for task', zodat iedere medewerker de stuurinformatie direct in zijn of haar proces kan gebruiken. Daarmee wordt het gebruik en de democratisering van Analytics een feit.

Mirjam Hulsebos is freelance journalist.