

Eindhovense wetenschappers slaan handen ineen voor Visual BPM

Plaatjes halen waarde uit data

Frans van der Geest

Een plaatje zegt meer dan duizend woorden. Dankzij onze natuurlijke, visuele geaardheid zijn we in staat complexe materie te begrijpen. Landkaartmakers passen de truc al jaren toe. Nu is het de beurt aan BPM-specialisten om bedrijfsprocessen inzichtelijker te maken met aantrekkelijke plaatjes.

In een presentatie naar de pers in 1976 deed de vermaarde Britse IT-consultant James Martin de voorspelling dat computerprocessors ooit zo snel zullen zijn dat we niet voldoende data voorhanden hebben om ze te voeden. Martin werd in het jaar daarop genomineerd voor de Pulitzer prijs voor zijn boek 'The Wired Society'. Het bleef bij die nominatie en achteraf is dat misschien maar goed ook, want inmiddels weten we dat het Engels data-modelleer-icoon er behoorlijk naast zat met zijn voorspelling. Ongekend groot is onze rekenkracht en tot in de haarvaten van de samenleving volop beschikbaar. Doch de databerg dijt uit. In 2007 kende de wereld naar schatting een opslagcapaciteit voor zowel analoge als digitale data voor in totaal bijna 295 miljard Gigabytes. Nu, vier jaar later praten we over 988 miljard Terabytes. Het einde is nog lang niet in zicht. Want alles dat eerst alleen op papier stond of op film- dan wel fotomateriaal, gaat het digitale archief in. En dan hebben we ook nog de onvoorspelbare toename van het mobiele dataverkeer in samenhang met allerlei vormen van social media. Nu al maken naar schatting 1,2 miljard mensen dagelijks gebruik van een mobiel apparaat om informatie vast te leggen. En praktisch al die handelingen of transacties leiden tot registraties en dus weer tot nieuwe data.

Een normaal mens beleeft slapeloze nachten bij de gedachte grip te moeten houden op die data-explosie. Mensen die wat slimmer zijn, slapen wellicht ook niet, maar gebruiken de wakkere uren om vernuftige systemen te bedenken voor het oplossingen van het vraagstuk. Daarom organiseren ze bij de TU-Eindhoven regelmatig zogeheten BPM Round Table bijeenkomsten, waar afgelopen maart het thema Visual Analytics centraal stond. Want willen we bij het beheren van bedrijfsprocessen iets halen uit de gigantische databerg, dan zullen we als mensen onze visuele eigenschappen moeten aanspreken. In zijn inleiding schetste

TU-hoogleraar prof. dr. ir. Wil van der Aalst de definitie van het raakvlak tussen Business Process Management en Visual Analytics. Hij spreekt van process mining met als doel event gebonden data te gebruiken voor het extraheren van procesgerelateerde informatie. Je wil als het ware automatisch een procesmodel kunnen herkennen bij het observeren van events die door informatiesystemen zijn gegenereerd. Door data en processen zichtbaar te maken, kunnen mensen er iets mee.

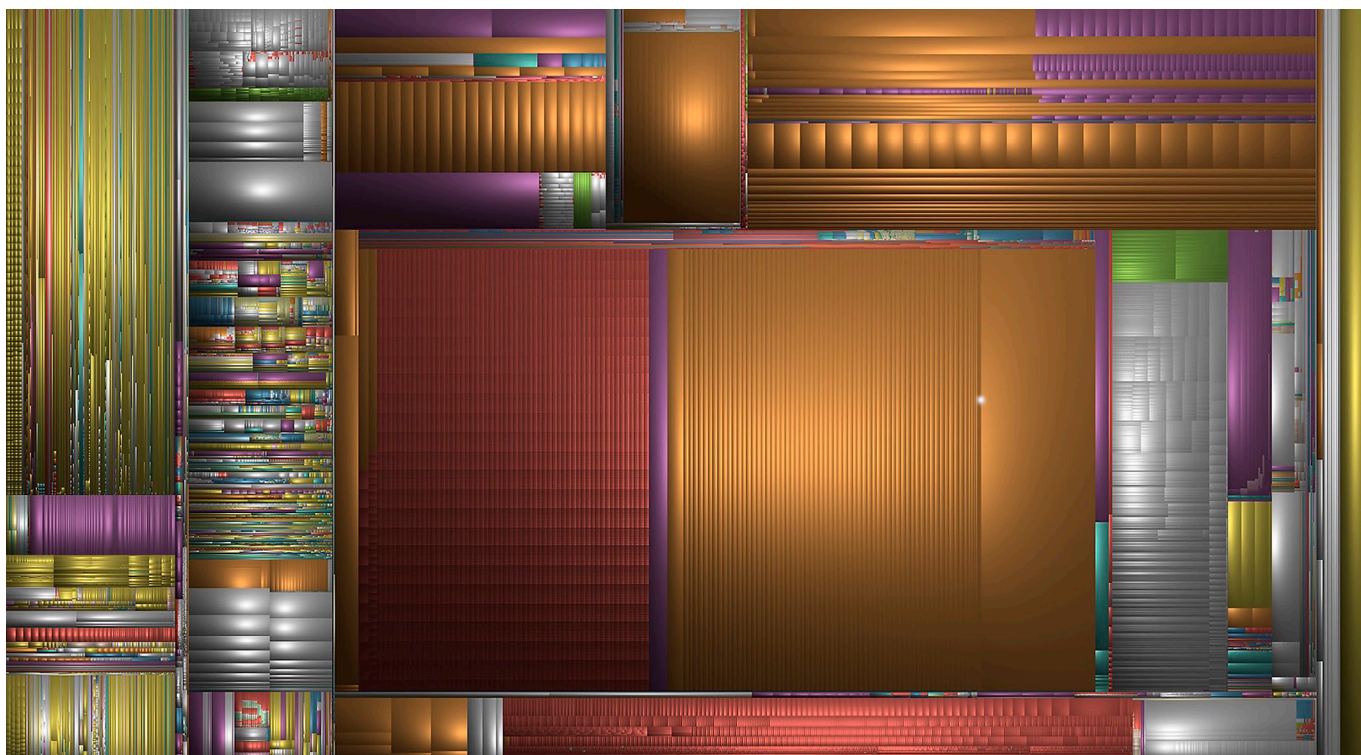
Wegenkaart versus stroomdiagram

Zo vormt een wegenkaart met al zijn nuances in kleuren en symbolen volgens prof. Van der Aalst in het algemeen een bruikbaar oriëntatiemiddel dan een abstract stroomdiagram zoals gebruikt in BPM-kringen. "We maken geen gebruik van kleur en we abstraheren niet zodat we geen dingen kunnen weglaten.

Plaatjes geven meer dan abstracte getallen een gevoel van urgentie

De plaats van symbolen is volledig willekeurig gekozen. En we doen alsof dit het enige juiste model is van de werkelijkheid. Van landkaarten weten we dat er meer weergaven van dezelfde werkelijkheid zijn. Als je met de fiets erop uit gaat, gebruik je een andere kaart dan wanneer je met de auto gaat reizen", aldus Van der Aalst.

In de Fuzzy miner in het process mining gereedschap ProM gaat het al beter. De visualisatiesoftware staat het gebruik van kleurbolletjes toe en biedt bovendien de mogelijkheid van in- en uitzoomen. Van der Aalst toont nog een voorbeeld uit ProM, de



Afbeelding 1: SequoiaView.

zogeheten social network miner, waarmee de onderlinge samenhang van sociale netwerken is weer te geven. Voor materiedeskundigen ongetwijfeld een nuttig instrument. Voor de leek levert het evenwel slechts een complex plaatje op dat waarschijnlijk eerder wordt beoordeeld als kunstwerk. Mensen die data of process mining tools hanteren, zoeken uiteraard naar heel andere verbanden dan toeristen die zich oriënteren in een bepaalde landstreek. Toch beschouwt Van der Aalst het als een uitdaging om procesmodellen net zo inzichtelijk te maken als een wegenkaart. Voor het halen van de fundamentele kennis van visualisatie behoeft hij niet ver te reizen. Aan de TU-Eindhoven treft hij een collega, prof. dr. ir. Jack van Wijk, die internationaal faam geniet op onderhavig terrein.

Waarom is de harde schijf vol?

Om zijn onderzoeksdomein te demonstreren toonde prof. Van Wijk een plaatje uit het in zijn vakgroep ontwikkelde SequoiaView, een inmiddels ruim tien jaar oud programma waarmee de bezetting van de harde schijf van een computer op heel duidelijke wijze is weer te geven. Iedereen die zich afvraagt waarom zijn harde schijf vol zit, doet er goed aan dit stukje universitaire huisvuil te downloaden (www.win.tue.nl/sequoiaview), want het functioneert een stuk beter dan Windows Explorer. Maar dat is er natuurlijk ook niet echt voor gemaakt. Ruim een miljoen mensen hebben inmiddels SequoiaView gedownload. Met rechthoeken worden de diverse bestanden in beeld gebracht. Hoe groter de rechthoek, des te groter het bestand. De kleur geeft aan om welk type bestand het gaat; zo staat paars voor een (ZIP) archiefbestand. Plaatjesbestanden zijn rood van

kleur en oranje staat voor video, allemaal Gigabytevreters dus. Maar omdat ook de structuur van de bestanden is weergegeven, wordt het bij verder inzoomen duidelijk dat de diverse kleine Service Pack files van Windows met elkaar ook behoorlijk wat ruimte innemen.

Information Visualization is de formele benaming van het vakgebied van prof. Van Wijk, die de toehoorders een definitie niet wil onthouden. Hij omschrijft het als de studie naar het gebruik van computerondersteunde, interactieve, visuele representaties van abstracte data om de weergave daarvan te versterken. De onderliggende actie is verrassend simpel: er is een abstracte dataset (tabel, grafiek of boomdiagram); die wordt via een systeem gevisualiseerd naar een beeld voor de gebruiker; die beoordeelt het beeld en interacteert met het systeem over zijn waarneming. Nu legt het Eindhovense onderzoeksteam zich vooral toe op softwarevisualisatie waarin diverse belangengroep acteren. Zo zijn er programmeurs die code inkloppen, architecten die systemen ontwerpen en managers die ontwikkelprojecten leiden. Hun vragen zijn volgens prof. Van Wijk nog een stuk waziger dan iemand die wil weten waarom zijn harde schijf vol zit. Dus alleen via een plaatje valt er iets te doorgronden, hetzij via een node link diagram, een boomdiagram of een meervoudige variabele weergave in de vorm van een plot van verspreide puntjes.

Wereld ingedeeld in drie dataklassen

De doorsnee vragen waarop het beeld antwoord moet geven, hebben betrekking op zaken als: is de structuur van een softwaresysteem netjes; valt die eventueel te verbeteren; is de implementatie verlopen volgens de vooraf vastgelegde architectuur;

Schone schijn of hard feit

'The Beauty of Data Visualization' luidde de prikkelende titel van een presentatie gegeven tijdens The Premier Business Leadership Series in Antwerpen, georganiseerd door SAS Institute. Deze pionier in analytische applicaties realiseert zich als geen ander dat informatie pas gaat leven wanneer 'droge' cijfers 'sappig' worden gepresenteerd. En of we nu topmanager, bedrijfsmedewerker of gewoon consument zijn, iedereen heeft een visueel hulpmiddel nodig om z'n verbeeldingskracht te activeren. Sommige SAS-gebruikers zweren bij de 'dashboards'.

David McCandless, de geestelijk vader van 'The Beauty of Data Visualization', is daarvan minder overtuigd. Hij heeft veel 'boardroom dashboards' onder ogen gekregen die het niveau van een matige PowerPoint presentatie nauwelijks ontstijgen.

De aanwezigheid op de SAS-conferentie schotelt hij andere koek voor. Met één simpel ingekleurd plaatje laat hij iedereen kennismaken met de dramatisch verslechterde visstand in onze zeeën en oceanen. Het is een feit, onweerlegbaar voor elke criticaster van ons natuurbeleid. Althans die zeggingskracht heeft het eenvoudige plaatje. Aan vervaardigen ervan ging veel onderzoek vooraf. Cijfermateriaal, online vergaard uit vele bronnen werd ingeklopt in een spreadsheet voor calculatie en ordening.

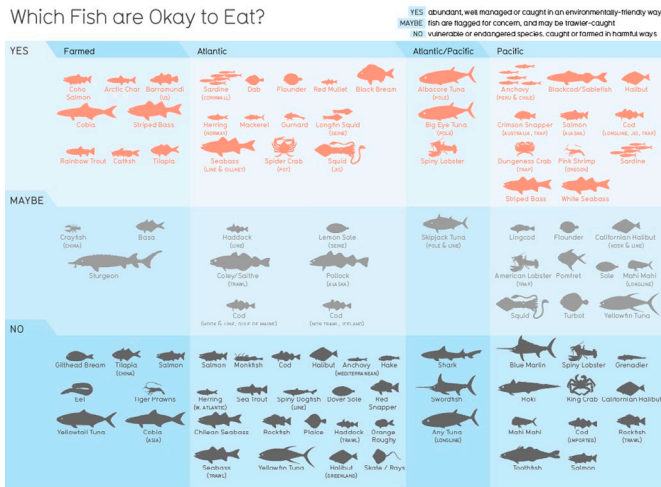
Pas daarna borrelt bij McCandless het idee op hoe het resultaat te visualiseren. "Je moet wel uitkijken om weergaven niet te veel te verfraaien. Beelden leiden gauw af van de werkelijke betekenis. Ik onderbouw mijn visuele conclusies altijd met de onderliggende databronnen", aldus de briljante Brit.

Dit streven naar waarheidsgehalte past in de professe van David McCandless. Hij is datajournalist. Hij produceert zijn kopij door louter grote hoeveelheden digitale informatie te raadplegen en door middel van data mining technieken en slimme filters, verbanden te leggen tussen de data-elementen. Het vak heeft hij zichzelf geleerd na eerst op de traditionele wijze journalistiek te hebben bedreven. Ook bij het toepassen van visualisatietechnieken is hij tot nu toe zelfstandig te werk gegaan en ligt het accent meer op intuïtie dan op het gebruik van wetenschappelijk onderbouwde methoden. Hij verwacht overigens wel dat de invloed daarvan gaat toenemen. McCandless noemt de City University in Londen als één van de instituten die in het fundamenteel onderzoek op het snijvlak van data mining en visualisatie het voortouw in zijn land gaat nemen.

waar zitten de zwakke schakels en – belangrijk voor managers – waarom heeft de ontwikkeling zoveel mankracht, tijd en dus geld gekost? "Op het moment dat je het probleem in een formule kunt vatten, heb je het in feite opgelost," zegt prof. Van Wijk. "Je laat er een programmaatje op los en dat berekent een metriek. Zo krijg je snel een lijstje met de top tien van de grootste problemen. Maar in vele gevallen is er niet zo maar een formule voorhanden. Vooral niet bij problemen waarvan men zich niet bewust is. Een plaatje biedt dan uitkomst met als uitdaging dat de belanghebbenden vanuit de verschillende betrokken disciplines daarop allemaal hun eigen kijk hebben. Iemand die met een *debugger* bezig is, kijkt meer naar het executieniveau en heeft behoefte aan een heel andere weergave. Toen ik met dit vak begon dacht ik dat de wereld netjes was opgedeeld in drie klassen van data. In werkelijkheid blijkt het een ingewikkelde mix. Over onze software kunnen we metrieken uitrekenen, we herkennen de hiërarchie met systemen, subsystemen en modulen en netwerkdata, die laten zien hoe de diverse componenten in het systeem elkaar aanroepen. Verder speelt ook tekst een rol, ofwel in de code zelf ofwel voor de identificatie van de diverse gebruikte componenten. De uitdaging is om van deze ingewikkelde combinatie een plaatje te maken, zodat afhankelijk van de vraag men meer of minder aandacht aan het ene dan wel het andere aspect kan besteden."

Hoewel zijn experimenten zich toespitsen op softwarevisualisatie, denkt Van Wijk dat dezelfde technieken ook toe te passen zijn in Business Process Management omgevingen. Zijn voorbeelden uit de softwareontwikkeling herbergen thema's die zich zonder meer ook laten projecteren tegen een meer procesmatige

Which Fish are Okay to Eat?



Afbeelding 2: Overzicht visstand, aan vervaardiging ging veel onderzoek vooraf.

achtergrond. Zo toont hij een matrixweergave van een grafiek ten behoeve van een incidentenrapportage. Voor een software architect die wilde weten of iedereen zich aan zijn ontwerp heeft gehouden, werd een boomdiagram gecombineerd met een grafiek ontwikkeld. Daarin zijn de begrenzinglijnen hiërarchisch gebundeld. Onderzoek toonde aan dat door aan de lijnen verschillende richtingen te geven de visuele weergave is te beïnvloeden. Je kunt met pijltjes op de lijnen werken of met lijnen die van licht naar donker verkleuren dan wel net andersom. Van groene naar rode lijnen levert bruikbare weergaven op, evenals bochtige lijnen en taps toelopende lijnen. Bij een vergelijking

met elkaar blijkt het bochtenwerk het minst interessante resultaat op te leveren. De taps toelopende varianten doen het heel goed, net als de van donker naar licht lopende lijnen.

Van laboratorium naar praktijk

Over smaken valt niet te twisten. Eigenlijk kan hetzelfde gezegd worden over visuele weergaven. Terwijl de één zweert bij kleuren en symbolen, geeft de ander de voorkeur aan saaie cijferreeksen. Demo's van Business Intelligence applicaties bevatten vaak fraai meters, die de indruk moeten wekken dat bedrijfsbestuurders net als in een auto zich laten leiden door de plaats van de naald op de achtergrond met de kleuren groen, oranje, rood. In de Business Intelligence praktijk hebben die meters zich als volledig onbruikbaar bewezen. En misschien worden hun equivalenten in de auto ook zelden door de bestuurder geraadpleegd. In de wonderde wereld van de visualisatietechnieken ervaart men dagelijks proefondervindelijk wat wel en wat niet werkt. En dat niet alleen vanuit het laboratorium. De TU-Eindhoven vormt de bakermat voor veel bedrijven, ontstaan uit wetenschappelijk onderzoek. Zo kon het gebeuren dat de SequoiaView software van prof. Van Wijk op een notariskantoor belandde. Daar is het visualisatiehulpmiddel niet ingezet om de vullinggraad van de harde schijf uit te beelden, maar om het rendement op de diverse dossiers inzichtelijk te maken. Voor de notaris een nuttig instrument om de prestaties van zijn kantoor te monitoren, voor de visualisatiewetenschappers het overtuigende bewijs dat hun kennis zich laat vertalen naar klinkende munt. Het bedrijf MagnaView zag als spin-off van de universiteit het levenslicht. Inmiddels is het een serieuze onderneming, die binnen en buiten Nederland software en bijbehorende diensten levert.

CEO is dr. Erik-Jan van der Linden. "Waarde uit data halen", aldus omschrijft hij heel beknopt de missie van zijn bedrijf. Daartoe wordt software ontwikkeld die breed toepasbaar is in vele vakgebieden en industrietakken. Met die software wil MagnaView de brug slaan naar de gebruikers die grote hoeveelheden data interactief willen interpreteren en vertalen naar zakelijk besluitvorming. Van der Linden: "Visuele weergaven laten hen de onderliggende data begrijpen. Daarnaast geven plaatjes meer dan abstracte getallen een gevoel van urgentie, vergemakkelijken ze het volgen van het analysetraject en staan ze professionals toe over hun bevindingen onderling te communiceren". De visualisatieoplossingen die zijn club tot nu toe heeft gemaakt richten zich op heel specifieke doelgroepen. Zo doen klinisch pathologen en microbiologen er hun voordeel mee evenals juristen en onderwijsmanagers. De specialistische kennis van deze gebruikers is als het ware geïnjecteerd in analysealgoritmen die vervolgens plaatjes doen ontstaan, waaruit de materiedeskundigen onmiddellijk conclusies kunnen trekken. De vorm van de visualisatie is volgens Van der Linden zeker niet generiek. Zo zijn er doelgroepen die helemaal geen kleur willen of andere contouren van cosmetische aard. Verrassend genoeg doen in sommige kringen tabellen het veel beter.

Procesdata in systemen

De praktijk blijkt dus weerbarstig. Dat geldt volgens Van der Linden ook voor het combineren van Visual Analytics met business processen. De mate waarin dat mogelijk is, hangt af van het feit of procesdata wel zijn vastgelegd in systemen. Zo bestaat er doorgaans weinig behoefte aan inzicht over de voortgang bij het organiseren van een verjaardagsfeestje. Het protocol zit domweg in ieders hoofd. Intuïtief doet iedereen wat er van hem of haar wordt verlangd. Ook een huisarts legt maar voor een heel klein deel de stappen van zijn werkproces vast, ook al bedienen ze zich steeds vaker van een computer tijdens de spreekuren met hun patiënten. Wie een bezwaarprocedure opstart, merkt al snel dat de actiestappen voor het overgrote deel zijn geprotocolleerd en in de procesindustrie zit alles in systemen. De volgorde van de processtappen is nauwkeurig vastgelegd. In de bestuurlijk administratieve wereld ligt dat anders. Daar is meestal sprake van transactiegeoriënteerde systemen. Die bevatten wel een schat aan procesdata. Als het systeem de processen aanstuurt, zal visualisatie de business analist in staat moeten stellen het proces te verbeteren door het systeem aan te kunnen passen. Bepaalt een materiedeskundige de stappen in een proces, dan richt de visualisatie zich puur op de operationele besluitvorming.

Frans van der Geest is freelance journalist.



BI-initiatieven mislukken nog veel vaker dan andere projecten. De BI-initiatieven moeten van de harde en de zachte kant komen. En als de harde kant van BI al praat met de zachte, spreken ze niet elkaars taal. Het boek BI-ware is een boek voor ICT'ers en voor managers en vertelt in gewoon Nederlands wat er allemaal fout kan gaan en wat daaraan kan worden gedaan. BI-ware bevat een bundeling van artikelen van Karien Verhagen en is een nieuwe uitgave in de reeks van DB/M Essays. De artikelen zijn gepubliceerd in de periode 2002 – 2006.

Wilt u weten hoe u Business Intelligence kunt laten slagen? Dan kunt u niet zonder deze uitgave! Ga snel naar www.array.nl en bestel BI-ware!

Deze uitgave is mogelijk gemaakt door:  