

Niet alleen verplichte kost voor specialisten

ITCARRIÈRE

Data scientist: een beroep met per



Albert Bogaard, partner van Ortec:
"Iedereen zou kennis van data analytics, Big Data en data science moeten hebben."



Pascal Lubbe van SAS:
"Bedrijven vinden de investering in tijd van hun medewerkers bezwaarlijker dan de cursuskosten."

Het is een van de meest gevraagde functies in de IT: data scientist, een persoon die in staat is om uit onafzienbare hoeveelheden data de juiste trends te halen, waar bedrijven snel op kunnen inspringen. Er wordt bijna jacht gemaakt op de schaarse mensen met deze vaardigheid. Maar verstand van data-analyse is niet alleen iets voor specialisten. Verschillende initiatieven springen in het gat dat is ontstaan.

door: RICHARD KEIJZER / R.KEIJZER@AUTOMATISERINGGIDS.NL

Het vak data scientist kent zijn oorsprong in de Verenigde Staten en daar zijn dan ook de eerste opleidingen tot stand gekomen. De rest van de wereld is nu bezig aan een inhaal-slag; ook Nederland heeft leergangen van hoog niveau. De vraag komt van het bedrijfsleven en ook van de overheid en academische instellingen. Overal bestaat de wens om tonnen ruwe gegevens om te zetten in bruikbare informatie. Eind 2014 kwam de Nationale Denktank tot de conclusie dat er in 2018 in Nederland alleen al een tekort is van 8000 data scientists. Dat de data scientist een gewild persoon is, blijkt wel uit het speciale programma dat venturecapitalbedrijf Greylock Partners heeft opgezet. Greylock, dat in het begin ook veel geld heeft gestoken in Facebook en LinkedIn, heeft een speciaal team van headhunters gevormd. Dat team

struikt het land af naar feitelijke en potentiële data scientists, om ze vervolgens een baan aan te bieden bij een van de bedrijven waarin Greylock heeft geïnvesteerd.

Niet onoplosbaar

Dat tekort is niet onoplosbaar, meent Albert Bogaard, partner van Ortec. "Ik vergelijk het wel eens met de Cobol-kraaiers van een jaar of dertig geleden. Daar waren er ook maar bitter weinig van. Toch denk ik dat je net zo gemakkelijk iemand tot data scientist kunt opleiden als tot Cobol-programmeur. Vergeet ook niet, een data scientist hoeft niet iemand te zijn die net van school af is. Ook mensen die een nieuwe uitdaging zoeken, kun je proberen om te scholen tot data scientist. Dat besef zie ik bij het UWV trouwens te weinig, ze zouden daar best wat alerter

spectief



Monique Rutten, strateeg bij de Data Driven Marketing Association: '...bedrijven helpen wat te doen met de goudmijn die de afgelopen jaren is ontstaan door het verzamelen van data'.

op mogen zijn. En managers zouden ook eens beter naar hun eigen personeel kunnen kijken. Zitten daar soms mensen tussen die aanleg hebben om te worden opgeleid tot data scientists?"

Bogaard stond aan de basis van één van de initiatieven die het gat tussen vraag en aanbod wil dichten. Bogaard zag de versnippering op het gebied van Big Data en data science met lede ogen aan en besloot daar wat aan te doen. In 2014 richtte hij daarom samen met Han van Dissel en Marc Salomon van de Universiteit van Amsterdam de Big Data Alliance op. De alliantie slaat een brug tussen de grote Nederlandse universiteiten en instellingen als het Centrum voor Wiskunde en Informatica (CWI), SURFsara, Amsterdam Data Science en het Data Science Center Eindhoven (DSC/e). Het gremium wordt voorgezeten door Alexander Rinnooy Kan.

Eén van de praktische resultaten van die samenwerking is The Analytics Academy, een school die is ontstaan uit een hechte samenwerking tussen het bedrijfsleven en de academische wereld. De leergangen zijn bestemd voor mensen uit diverse disciplines. "Eigenlijk zou iedereen kennis van data analytics, Big Data en data science moeten hebben", verklaart Bogaard de breedte van het opleidingsaanbod. "Ga maar na, managers lezen van alles en nog wat over het onderwerp terwijl ze nauwelijks weten waar het werkelijk over gaat. De boardroom hoeft niet weken op cursus, voor executives is een korte leergang van een paar uur tot hooguit een dag of twee meer dan voldoende. Ze leren dan de basisprincipes, zodat ze een begrip krijgen voor de data die in hun bedrijf aanwezig is en wat mensen daar allemaal mee kunnen doen."

Technici en specialisten hebben niet genoeg aan een paar dagen cursus. Voor hen zijn uitgebreidere leergangen opgezet. De echte datadeskundigen kunnen bijvoorbeeld gebruik maken van twee cursussen, die allebei zo'n 7 weken in beslag nemen. Het gaat om de leergangen Data Scientist en Boundary Spanner. "Een data scientist gaat de diepte in, kijkt onder de motorkap en leert werken met allerlei technieken en programmeertalen. Een boundary spanner maakt ook gebruik van de techniek, maar kijkt veel meer naar logische verbanden in de data. In dit deel van de cursus wordt de creativiteit ook geprikkeld om de cursisten goede resultaten te laten bereiken."

Creativiteit is voor een boundary spanner van groot belang. Het stelt hem in staat om geavanceerde koppelingen te maken tussen gegevens. Daarbij is een controle op de bedachte analysetechnieken en algoritmen zeer belangrijk. Een kleine fout kan namelijk grote gevolgen hebben. Bogaard: "We kennen allemaal wel de Excel-foutjes, een verkeerd getal in een cel van het rekenblad, waardoor de resultaten soms compleet uit de richting kunnen zitten. Bij data analytics heb je hetzelfde gevaar, maar dan op een zeer grote schaal. Die controle, daar heb je een mens voor nodig, een machine kan dat niet."

Praktisch gebruik

Het analyseren van grote hoeveelheden gegevens kan direct leuke resultaten opleveren. Bogaard: "Het hoeft niet direct om veel geld te gaan, gemak en tevredenheid zijn ook goede resultaten. Stel, dat je alle gegevens van je medewerkers in een model stopt, je kunt ze dan op bepaalde tijden naar huis laten gaan zodat ze zonder opstoppen op de oprit naar de snelweg kunnen doorrijden. Of denk aan het bedrijfsrestaurant, waar je mensen gefaseerd naar toe kunt laten gaan zodat ze niet in de rij hoeven te staan en meteen ook een tafel hebben."

Het beste is snel een toepassing in te voeren. "Men moet er op een agile manier mee gaan werken, zodat men meteen resultaat krijgt. Tijdens de hele opleiding hechten we daarom erg agile development. Snel met de materie aan de slag, uitproberen en kijken wat voor resultaten je er uit kunt halen. Doe je dat niet en zit je een paar maanden gewoon in de schoolbanken, dan is de wereld alweer veranderd als je klaar bent. De geleerde kennis is vrijwel niet meer toepasbaar in de nieuwe omgeving. Datzelfde geldt voor mensen die een app voor data analytics gaan maken. Die kunnen zich ook niet een jaar opsluiten in een laboratorium."

Amsterdam en Eindhoven zijn niet de enige universiteiten die het gat proberen te vullen. Tilburg University krijgt bijvoorbeeld een MSC-specialisatie Data Science, waarmee men wil inspelen op de stijgende behoefte aan data scientists door deelnemers aan de opleiding te doorkneden in het analyseren van megahoeveelheden data, het herkennen van patronen en de koppeling van de resultaten naar de echte wereld.

Marketeers

Initiatieven om meer data scientists op te leiden komen niet alleen uit universitaire hoek. De Data Driven Marketing Association (DDMA) heeft met dat doel voor ogen de Data Academie opgericht. Die academie moet bedrijven helpen 'wat te doen met de goudmijn die de afgelopen jaren is ontstaan door het verzamelen van data'. "Het kan gaan om Big Data", zegt Monique Rutten, strateeg bij de Data Driven Marketing Association (DDMA), "maar dat hoeft niet. Small Data kan net zo waardevol zijn. De vraag is echter, wie die waarde uit de data gaat halen en wat voor gereedschappen hij of zij daarbij gaat gebruiken. Kennis is een voorwaarde om goed met data overweg te kunnen. Via de Data Academie geven we toepasselijke cursussen."

De Data Academie richt zich op de marketeer. "Er zijn verschillende types, die allemaal een eigen opleiding nodig hebben. Daar zorgen we dan ook voor", zegt Rutten. "Van oudsher kennen we de klassieke marketeer die aan de hand van vuistregels, productkennis en branding zijn strategie uitstippelt. Dan is er de databewuste marketeer die weet wat data voor marketing kan doen, maar zijn hart maakt geen sprong bij Excel. Dit gebeurt wel bij de dataminende marketeer. Hij neemt de beschikbare data uit en te na tot zich, laat er analyses op los en laat de uitkomst daarvan zijn strategie bepalen."

Alle type marketeers moeten leren wat data voor hen kan doen. "Ze hoeven overigens geen specialisten te zijn die trends haarfijn kunnen halen uit grote hoeveelheden data. Maar ze moeten wel weten welke vragen ze moeten stellen om het beste begrip te krijgen van de data. De data kan hen vertellen welke touch points er zijn met hun

(potentiële) klant. Deze informatie maakt het weer mogelijk een customer journey in kaart te brengen die waardevol is bij de werving van nieuwe klanten”, meent Rutten.

Kleinschalig

De opleidingen die door de Data Academie worden gegeven zijn kleinschalig opgezet. Rutten: “Je moet denken aan klasjes waar ongeveer acht mensen in zitten.” Wanneer een cursus in huis wordt gegeven hangt het aantal deelnemers samen met de grootte van het marketingteam. De cursussen worden op een aantal niveaus gegeven. Voor de allround marketeer gaat het er om dat ze gevoel krijgen voor data. De cursussen voor de data-experts zijn veel specialistischer van aard. Denk aan voorspelmodellen en regressie.”

“In het afgelopen jaar hebben we gemerkt dat de cursisten een analytisch vermogen hebben en nieuwsgierig zijn. Ideale ingrediënten voor een moderne marketeer. Omdat alle type marketeers in de klas kunnen zitten, bekijkt de cursusleider wat voor mensen hij voor zich heeft. Een pure analist benaderen we heel anders dan een brand marketeer. De eerste zal meteen op een stapel data duiken, de tweede zal afwachten en kijken wat de kennis uit de data voor zijn merk kan betekenen. We hebben een tiental cursusconcepten ontwikkeld. Omdat we de kennisbehoefte precies willen laten aansluiten op de geboden informatie, houden we na afloop iedere cursus opnieuw tegen het licht. Mocht dit nuttig zijn, dan passen we inhoud of vorm aan tot er een vernieuwd concept ontstaat”, zegt Rutten.

Niet te langdurig

Belangrijk is ook, dat de marketeers niet te lang uit hun werk worden gehaald. Daarom zijn de cursussen van de Data Academie niet te lang. “In één, hooguit twee dagen moeten we de lesstof kunnen overdragen en moet het zaadje worden geplant”, zegt Rutten, “want het gaat om de praktijk en daar moet de verworven kennis worden toegepast.”

Mensen die een van de cursussen van de Data Academie hebben gevolgd, krijgen na afloop een meldingstekst, die ze op hun LinkedIn-profiel kunnen opnemen. Iedereen kan dan zien welke kennis ze hebben verworven. Rutten: “We hebben overwogen om onze cursussen te laten certificeren, maar dat is een zeer langdurig proces. Het duurt vaak zo lang, dat de cursus in kwestie alweer geactualiseerd is op het moment dat de certificatie rond is. Dat aspect hebben we dan ook laten varen.” Om van een cursus een succes te maken, dient er een expert voor de klas te staan. De vraag is, hoe je die experts vindt? Rutten: “We hebben mensen nodig met de beste data-expertise, om het mooi te zeggen. Die zitten overal en de kunst is om ze te vinden. Een groot aantal zit al in het netwerk dat we als brancheorganisatie hebben en gaandeweg wordt dat aantal uitgebreid met nieuwe mensen. Opleiders komen bijvoorbeeld uit de universitaire hoek, maar ook van consultants en adviesbureaus. En voor goede kennis betaal je. Dit vertaalt zich in de marktconforme opleidingsprijzen. De academie is non profit dus het mooie is dat de inkomsten die onder aan de streep overblijven, weer terugvloeien naar de ontwikkeling van de Data Academie. Dit maakt het een soort voor-en-door-de-branche initiatief.”

Bedrijfs cursussen

Er zijn ook bedrijven die zelf cursussen organiseren. Zo heeft SAS begin dit jaar het startschot gegeven voor een nieuwe opleiding voor data scientists. Op 9 januari trad de eerste lichter aan. De opleiding neemt maar liefst 13 weken in beslag. In de cursus worden de cursisten via een combinatie van klassikale lessen en praktijkoefeningen de fijne kneepjes van het vak bijgebracht. “Er zijn veel snellere cursussen, maar wij denken dat je wel meer nodig hebt dan een paar uren of dagen om het vak meester te worden”, zegt Pascal Lubbe van SAS. De opleiding gaat eerst van start in de VS, later zal hij ook in Europa worden gegeven. “We noemen het de Academy for Data Science en de aftrap is begin januari gegeven”, zegt Lubbe.

De deelnemers aan de NL Academy for Data Science (ADS), die in september van start gaat, gaan aan de slag met gegevens uit de praktijk, met oplopende moeilijkheidsgraad. “Daarbij is het bedoeling dat ze meteen al de kans krijgen om fouten te maken, daar leren ze het meest van”, zegt Gerard Hellwig, als externe deskundige betrokken bij de ADS. “Denk bijvoorbeeld aan het inzetten van veel te veel computerpower om een probleem op te lossen. Als je slim nadenkt, kun je ook met minder resources toe. Dat maken we de cursisten meteen duidelijk”, stelt Hellwig.



James Kobielus, bigdata-evangelist bij IBM:
“Noem data scientists maar rustig tovenaars die in korte tijd waardevolle inzichten kunnen destilleren uit hopen data.”

WAT MOET EEN DATA SCIENTIST KUNNEN?

“De hedendaagse maatschappij wordt wel aangeduid met de term ‘Insight Economy’, ofwel een economie die berust op zichten. Inzichten die worden gegeven door mensen met de titel data scientist. Zij zijn in staat om grote bergen gegevens om te zetten in nuttige informatie”, zegt James Kobielus, bigdata-evangelist bij IBM. Dat roept meteen de vraag op wat een data scientist dan precies is. “Noem ze maar rustig tovenaars die in korte tijd waardevolle inzichten kunnen destilleren uit hopen data. Ze zijn slim en vindingrijk en bovendien weten ze hun kunde goed te verkopen. Dat lukt uiteraard prima in een markt waar een grote schaarste heerst”, aldus Kobielus.

Het vak van data scientist heeft een steile en lange leercurve en veel mensen laten zich afschrikken door het feit dat je veel kennis moet hebben van wiskunde en statistiek. En dat is dan nog maar de basis, de data scientist zal moeten bepalen welke variabelen uit de hoop Big Data het best geschikt zijn om er zakelijke inzichten door te verkrijgen.

En een goed inzicht ontstaat pas, wanneer de betreffende data scientist ook de juiste inzichten heeft en het kaf van het koren kan scheiden. Kobielus:

“Het is moeilijk om een exacte definitie te geven van de data scientist, maar het volgende lijstje eigenschappen past heel goed op die persoon:

- nieuwsgierig
- intelligent
- doorkneet in statistiek
- vermogen om te onderzoeken
- wetenschappelijk handelen
- en vooral sceptisch zijn.

De data scientist moet ook nog een goede communicatievaardigheid hebben, want hij zal de keuzes die zijn gemaakt moeten kunnen uitleggen aan anderen. Dan mag hij niet iets stamelen van ‘het leek mij het beste’. Nee, er zal duidelijk uitgelegd moeten worden waarom bepaalde waarden wel en waarom andere niet zijn meegenomen in de berekeningen.”

SAS levert de benodigde resources zoals opslag en verwerking vanuit de eigen cloud. Critici menen, dat deze insteek te beperkt is omdat de cursisten alleen leren werken met de software die bij SAS in huis is ontwikkeld. Later dit jaar zal SAS gaan samenwerken met de academische wereld, zodat het bedrijfstempel minder prominent op de leergang staat. De leergang bestaat uit twee delen die samen 20.000 euro kosten. “We merken nu al dat bedrijven de investering in tijd van hun medewerkers die naar de cursus gaan dan bezwaarlijker vinden dan de kosten. Na afloop van de cursus zijn de deelnemers – als ze slagen – gecertificeerd als Big Data Professional en data scientist”, zegt Lubbe.

