

Hulp bij spreadsheets

Voor professionalisering van Excel-gebruik moeten bedrijven eens bij hun IT-collega's kijken

Excel fungeert vaak als smeerolie tussen systemen: daar waar snel een flexibel model moet worden opgesteld, wordt Excel gebruikt. Dat heeft veel voordelen, zeggen **Arie van Deursen, Martin Pinzger en Felienne Hermans**, maar het gevolg is wel dat er in bedrijven honderden of misschien wel duizenden Excel-bestanden rondzwerven. Kleine foutjes leiden dan snel tot grote missers.

Spreadsheets worden veel gebruikt in het bedrijfsleven. Onderzoekers schatten dat 80 procent van de bedrijven in Europa gebruikmaakt van Excel voor financiële rapportages en dat maar liefst 90 procent van de analisten forecastingmodellen opstelt in Excel. Niet alleen de omvang van het gebruik van Excel is groot, ook de impact die spreadsheets hebben is enorm. Meer dan de helft van de spreadsheets wordt gebruikt als ondersteuning bij het nemen van keybeslissingen binnen een bedrijf. Maar een foutje in Excel is snel gemaakt: een formule klopt niet helemaal, data worden op de verkeerde plaats in een spreadsheet ingevoerd of een Excel-bestand blijkt te verwijzen naar een oude versie van brondata. Deze kleine foutjes leiden op regelmatige basis tot grote missers. Zo zorgden verkeerd geplaatste data er in 2005 voor dat Kodak zijn verlies met 9 miljoen dollar naar beneden moest bijstellen. En in 2006 werd het Amerikaanse gasbedrijf Dominion voor een

geschatte 200 miljoen dollar aangeklaagd. Waarom? Doordat ze een oude versie van een Excel-bestand hadden gebruikt om hun gasreserves te schatten, was de prijs van gas razendsnel gedaald. Toen bekend werd dat ze door een fout in Excel hun gasreserves flink overschat hadden, werden ze door concurrenten beschuldigd van inkomstenderving.

Een foutje in Excel is snel gemaakt

Een van de oorzaken van de problemen rond Excel is dat het niet altijd duidelijk is voor welke taken Excel gebruikt wordt en op welke manier die taken dan worden uitgevoerd. Excel fungeert namelijk vaak als de smeerolie tussen alle systemen: daar waar maatwerk net even tekortschiet of er snel een flexibel model moet worden opgesteld, wordt Excel gebruikt. Een voordeel hiervan is dat veel medewerkers

Excel goed beheersen en dus makkelijk en snel een bestand in elkaar zetten dat goed werkt. Een ander voordeel is dat een Excel-bestand makkelijk te delen is via een e-mailbijlage. Rapportages die uit softwarepakketten komen, worden vaak eerst naar Excel geëxporteerd, om op die manier rondgestuurd te kunnen worden. Ook heel makkelijk, maar deze voordelen zorgen ervoor dat er in bedrijven honderden of misschien wel duizenden Excel-bestanden rondzwerven. De situatie wordt nog eens bemoeilijkt door het feit dat er binnen een bedrijf soms verschillende versies van een Excel-bestand circuleren. De maker van een Excel-bestand stuurt een rapportage rond en iedereen die deze rapportage gebruikt, heeft dan zelf een versie op zijn pc staan, die hij zelf kan aanpassen zonder dat de andere gebruikers van het rapport dat weten. Zo ontstaan er na verloop van tijd veel versies van een bestand en is het niet altijd duidelijk wat de juiste of meest recente is. Ook is het vaak nergens

vastgelegd wie de eigenaar, dus de verantwoordelijke, van een bepaald bestand is.

Wat kan een bedrijf doen om de bovengenoemde 'foutjes' te voorkomen? Zoals gezegd worden er vaak bedrijfskritische berekeningen gemaakt in spreadsheets en is Excel in die zin soms net zo belangrijk als de IT binnen een bedrijf. De berekeningen die in Excel gemaakt worden, kunnen qua impact zeker onder de noemer 'ontwikkeling' worden geschaard. Ook de problemen rond spreadsheets lijken op de problematiek rond professionele softwareontwikkeling. Binnen softwareontwikkeling speelt versiebeheer een grote rol en ook daar wordt vaak de vraag gesteld hoe gedeeltes van de code, bijvoorbeeld klassen of libraries, samenhangen en waar variabelen worden gebruikt. Diverse tools en methoden die in de software al jaren gebruikelijk zijn, zoals het bijhouden van 'bugs', het geautomatiseerd en systematisch testen van software, zouden ingezet kunnen en moeten worden, met minder fouten en een beter begrip van Excel tot gevolg.

Internationaal spreadsheetonderzoek

De problemen rond spreadsheets zijn aanleiding voor veel onderzoek. Allereerst zijn er papers die de huidige problemen en hun oorzaken beschrijven. In 2001 publiceerde Barry Boehm, de uitvinder van COCOMO, samen met Victor R. Basili een lijst van de toptienproblemen die aangepakt zouden moeten worden om de kwaliteit van software in het algemeen te verbeteren. Op nummer tien plaatsten zij spreadsheets, waarbij ze opmerkten dat 40 tot 50 procent van de 'eindgebruikersprogramma's' fouten bevat terwijl de makers ervan overtuigd zijn dat de programma's foutvrij zijn. De European Spreadsheet Risk Group houdt op zijn website (www.eusprig.org) een lijst bij van 'horror stories' over fouten in Excel en hun oorzaken. Hierin zijn ook de verhalen over Kodak en Dominion terug te vinden. Daarnaast zijn er ook gebieden van onderzoek die zich richten op het verbeteren van de stand van

zaken. De eerdergenoemde Eusprig heeft daartoe vaakvoorkomende spreadsheetfouten gecategoriseerd en op basis hiervan een aantal best practices opgesteld voor spreadsheetgebruik, waaronder testen, documenteren en designen. Ook hier komen de uit de IT bekende concepten weer naar boven. Testen in het spreadsheetdomein kan bijvoorbeeld gedaan worden door onder aan ieder werkblad een aantal if-formules op te nemen met simpele controles, zoals waarin getest wordt of de som van alle percentages wel honderd is. Met conditional formatting kun je er dan voor zorgen dat je gewaarschuwd wordt als dit niet het geval is. Het klinkt wat omslachtig, maar als deze tests eenmaal geschreven zijn, kunnen ze de hele levensduur van de spreadsheet (soms meer dan tien jaar) helpen om fouten te voorkomen. Recentelijk beschreven Bill Bekenn en Ray Hooper

hoe diverse ontwerpprincipes uit Toyota's lean manufacturing toegepast kunnen worden in het ontwerpen van spreadsheets. Het principe van deze methodologie is dat richtlijnen simpel moeten zijn, omdat ze anders niet lang gebruikt zullen worden. De spreadsheetspecifieke richtlijnen die in dit artikel worden beschreven, zijn dan ook simpel van aard: hou data en formules gescheiden in aparte blokken, die verschillend gekleurd worden, zodat het duidelijk is wat voor operaties er waar in het bestand voorkomen. Ook wordt geadviseerd om gebruik te maken van sommaties over hele kolommen - SUM(A) - in plaats van over specifieke ranges - SUM(A1:A17) - om het risico van vergeten cellen te verminderen. Wetenschappelijk onderzoek levert dus diverse interessante aanknopingspunten om spreadsheetgebruik in de praktijk naar een hoger niveau te tillen.

Stappenplan

In samenwerking met het bedrijfsleven verricht de TU Delft onderzoek naar de problematiek rond spreadsheets. Zo bestudeert zij onder meer hoe spreadsheets omgezet kunnen worden in (UML)-klassediagrammen of in dataflowdiagrammen, om het werken met spreadsheets te ondersteunen. Ook wordt gekeken naar de problematiek rond Excel in de praktijk, onder meer bij een grote Nederlandse investeringsmaatschappij. In een dergelijke praktijkcase worden de volgende stappen gezet:

■ Probleemanalyse

In deze fase wordt door middel van interviews vastgesteld waar zich knelpunten bevinden. In de concrete casus bij de investeringsmaatschappij bleken deze te zitten in zogenaamde overdrachtsituaties, onder meer naar andere collega's, auditors of de IT-afdeling.

■ Oplossingsdefinitie

In deze fase wordt onderzocht hoe de problemen het beste verholpen kunnen worden. In het concrete geval waren met name gebrek aan inzicht in afhankelijkheden tussen werkbladen en (blokken met) formules problematisch. Een oplossing bleek het visualiseren van een Excel-bestand als een gelaagd dataflowdiagram. De bovenste laag geeft daarbij een Excel-bestand weer, de volgende laag weerspiegelt het niveau van werkbladen, en nog dieper liggen de individuele cellen (zie figuren 1 en 2).

■ Toolontwikkeling

In deze fase wordt de in het onderzoek ontwikkelde toolinfrastructuur verrijkt met de gevonden oplossingen. In dit geval werd de tool uitgebreid met de mogelijkheid om een gelaagde graaf af te

leiden uit spreadsheets en deze interactief te visualiseren middels de Directed Graph Markup Language.

■ Evaluatie

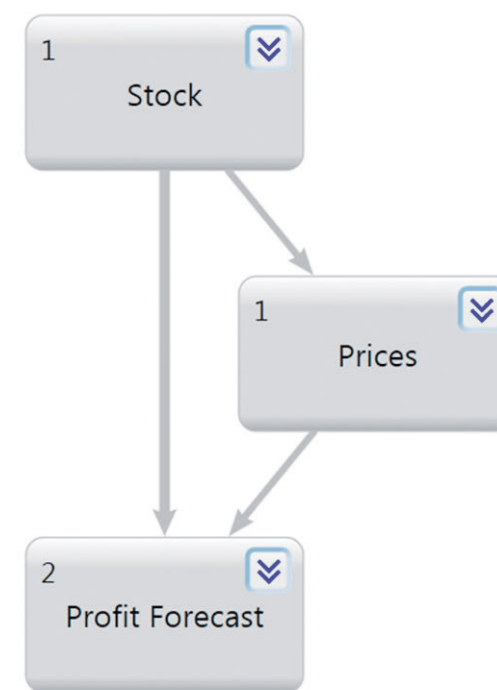
Ten slotte wordt gekeken in hoeverre de ontwikkelde aanpak de oorspronkelijke problemen oplost. Hiertoe werd het visualisatiegereedschap ingezet in negen concrete overdrachtsituaties. Bijna alle deelnemers (80 procent) vonden dat een dergelijke tool hun werkzaamheden makkelijker zou kunnen maken. De visualisatie bleek goed te helpen bij het uitleggen van het 'verhaal achter' de spreadsheet, bij het doorgronden van verwachte en onverwachte afhankelijkheden en bij het in detail begrijpen van specifieke formules.

Vaak is nergens vastgelegd wie de eigenaar van een bestand is

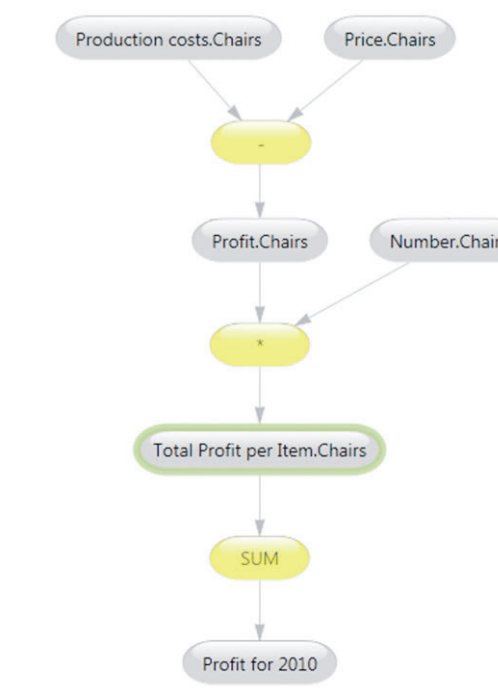
De algemene conclusie is dat bedrijven qua professionalisering van hun Excel-gebruik best eens wat mogen spieken bij de collega's van IT. Zowel de methoden die daar al jaren in gebruik zijn, zoals pair programming, peer reviews, als de tooling: versiebeheer, automatische analyse en visualisatie, zouden Excel naar een hoger niveau kunnen brengen.

Arie van Deursen is hoogleraar Software Engineering en hoofd van de Software Engineering Research Group van de TU Delft. Martin Pinzger is universitair docent aan de TU Delft en gespecialiseerd in softwarevisualisatie en -evolutie. Felienne Hermans verricht promotieonderzoek op het gebied van end-user programming. De auteurs zijn oprichters van Infotron (www.infotron.nl), een bedrijf gericht op het professionaliseren van het gebruik van spreadsheets.

Voor reacties en nieuwe bijdragen van deskundigen: Henk Ester (h.ester@sdu.nl), (070) 378 03 97.



Figuur 1: Spreadsheetvisualisatie op het niveau van tabbladen



Figuur 2: Spreadsheetvisualisatie op formulerniveau

ADVERTENTIE

TelecomExpo

Telecommunicatie, Netwerkinfrastructuur en Datacentertechnologie

Woensdag 8 en donderdag 9 december 2010 ■ Jaarbeurs Utrecht ■ Hal 9



GRATIS TOEGANG
BIJ VOORREGISTRATIE OP WWW.TELECOMEXPO.NL

Cloud Computing Congres

Gelijktijdig met TelecomExpo vindt het Cloud Computing Congres plaats, in het Media Plaza Utrecht. Het congres bestaat uit vier verschillende onderdelen. Het Telecom Management Association Jaarcongres maakt deel uit van het programma. U kunt zich voor ieder onderdeel apart aanmelden op www.cloudexpo.nl. U betaalt € 75,- per persoon per dagdeel voor deelname, abonnees betalen slechts € 50,- per persoon per dagdeel (prijs is excl. BTW). Dit is inclusief lunch/warm buffet op de beursvloer, koffie, thee en een uitrijkaart. Combineer uw beursbezoek met uw deelname aan het congres.

Organisatie: **Array PUBLICATIONS** **Telecom magazine** Mediapartner: **Automatisering Gids** Hoofdsponsors: **Aastra** **BT** **CISCO** **Microsoft**