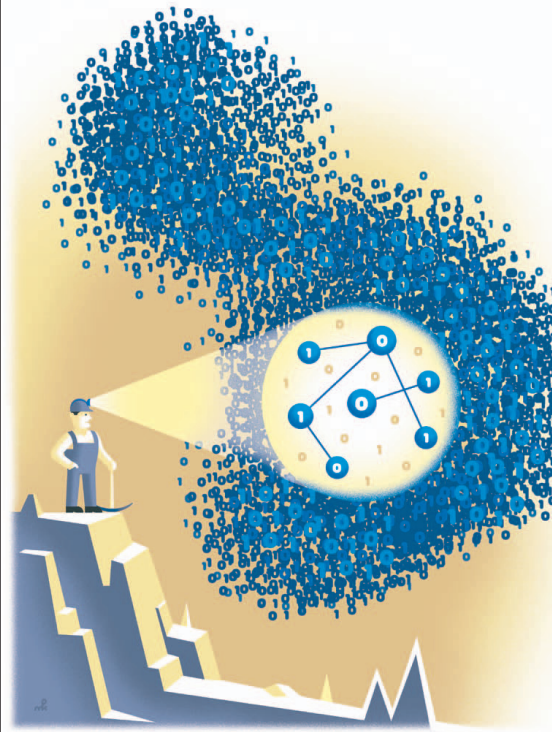


Het doel is data omzetten in betere processen en systemen

# PROCESS MINING IS ONTBREKENDE SCHAKEL

**Data worden verzameld maar niet optimaal gebruikt. De bestaande analysegereedschappen schieten tekort. Process-miningtechnieken, zegt Wil van der Aalst, zijn wel in staat modellen te genereren, waardoor data omgezet kunnen worden in nieuwe inzichten en betrouwbare antwoorden.**

door: WIL VAN DER AALST beeld: MARC KOLLE



## LANDKAARTEN VAN PROCESSEN

Process mining laat zien dat procesmodellen een soortgelijke rol kunnen vervullen als landkaarten. In beide gevallen gaat het om een abstractie van de werkelijkheid. Deze vergelijking laat meteen zien dat er voor een proces meerdere modellen mogelijk zijn, net zoals er vaak verschillende landkaarten van hetzelfde gebied zijn. Een automobilist heeft bijvoorbeeld een andere kaart nodig dan iemand die per metro reist. Welke landkaart het meest geschikt is hangt af van de vraag die beantwoord moet worden. Ook is het van groot belang actuele informatie te projecteren op de automatisch gegenereerde procesmodellen. Het kan bijvoorbeeld uitermate nuttig zijn de 'files' in een organisatie of systeem zichtbaar te maken.

**O**p dit moment staat 'Big Data' volop in de schijnwerpers. Het is voor iedereen duidelijk dat de exponentiële groei van gedigitaliseerde data en rekenkracht grote invloed zal hebben op onze maatschappij. Alleen organisaties die slim gebruik maken van de stortvloed aan gegevens over het gebruik van producten en diensten zullen overleven. Helaas ligt de nadruk van Big Data-initiatieven vaak op het genereren en opslaan van enorme hoeveelheden data in plaats van de analyse ervan. Als we een informatiesysteem zien als een auto en data als de brandstof die het systeem nodig heeft om te functioneren, dan wordt het probleem duidelijk. Het lijkt erop dat vele partijen bezig zijn auto's te bouwen die zoveel mogelijk brandstof gebruiken en een zo groot mogelijk tank aan boord hebben. De echte uitdaging is echter auto's te bouwen die snel en efficiënt zijn, dat wil zeggen het omzetten van data in betere processen en systemen. Hier schieten traditionele aanpakken gericht op data warehouses en simplistische dashboards tekort. Data worden verzameld maar niet optimaal gebruikt. Organisaties weten dat waardevolle antwoorden verborgen zitten in hun informatiesystemen, maar bestaande analysegereedschappen zijn niet in staat deze boven water te halen. Process-miningtechnieken zijn wel in staat modellen te genereren waardoor data omgezet kunnen worden in nieuwe inzichten en betrouwbare antwoorden. Bovendien zijn er steeds betere gereedschappen beschikbaar om dit snel en efficiënt te doen.

### Ontbrekende schakel

Steeds meer gebeurtenissen ('events') worden geregistreerd en opgeslagen in IT-systemen en dit gebeurt op allerlei niveaus. Denk bijvoorbeeld aan een ziekenhuisinformatiesysteem. Alle afspraken, onderzoeken, onderzoeksresultaten (bloedproeven, röntgenfoto's et cetera) en betalingen behorende bij een patiënt worden vastgelegd. Ook op het niveau van

## Software

Het open-source-gereedschap ProM ([www.processmining.org](http://www.processmining.org)) is de de facto standaard op het gebied van process mining in de wetenschappelijke wereld. Onderzoeksgroepen in Nederland, Duitsland, Frankrijk, China, Spanje, Italië, Rusland, Amerika en Australië hebben plug-ins voor ProM ontwikkeld. Het is voor iedereen beschikbaar en biedt in vergelijking met commerciële gereedschappen enorm veel functionaliteit. Binnen het open-source-gereedschap zijn er meer dan 600 plug-ins beschikbaar die het gehele process-miningspectrum afdekken. ProM is echter vooral bedoeld voor process-miningexperts en minder geschikt voor eindgebruikers. Commerciële gereedschappen zoals Disco (Fluxicon), Perceptive Process Mining (eerder Futura Reflect), ARIS Process Performance Manager (Software AG), QPR ProcessAnalyzer, Interstage Process Discovery (Fujitsu), Discovery Analyst (StereoLOGIC) en XMAAnalyzer (XMPro) maken process mining geschikt voor een groter publiek. Deze gereedschappen kunnen door eindgebruikers gebruikt worden om de onderliggende processen te ontdekken en knelpunten op te sporen. In vergelijking met ProM schieten ze echter te kort als het gaat om conformance checking en predictive analytics. Gereedschappen als Disco kunnen direct CSV-files inlezen en omzetten in XES (eXtensible Event Stream) of MXML (Mining XML)-files. De laatste twee standaardformaten zijn toegesneden op process mining en maken het delen van event-data eenvoudig.

## EVENT LOGS ZIJN HET UITGANGSPUNT



**Wil van der Aalst**  
([www.vdaalst.com](http://www.vdaalst.com)) is universiteitshoogleraar aan de Technische Universiteit Eindhoven (TU/e). Daarnaast heeft hij deeltijdastellingen bij Queensland University of Technology (QUT) in Australië en de National

Research University Higher School of Economics in Moskou. Hij is ook de trekker van de IEEE Task Force on Process Mining.

individuele apparaten worden events geregistreerd. Het maken van een röntgenfoto resulteert in een lawine van events die geanalyseerd worden door de leverancier die wil weten hoe het röntgenapparaat echt gebruikt wordt en wanneer en waarom het apparaat soms stukgaat. Het 'Internet of Things' (IoT) waarbij ook eenvoudige apparaten (van koelkast tot kopieerapparaat) gebeurtenissen registreren en delen, is steeds meer een realiteit.

Klassieke technieken voor procesanalyse en herontwerp maken geen gebruik van de grote hoeveelheden data die vandaag de dag beschikbaar zijn. De nadruk ligt op het handmatig maken van processchema's en de analyse van modellen in plaats van data. Omgekeerd schieten traditionele data-mining- en BI-(Business Intelligence)-technieken tekort vanwege de afwezigheid van een duidelijke procesoriëntatie. Eenvoudige dashboards en beslismomen kijken vaak naar geaggregeerde data of een enkel punt in het totale proces. Process mining is de ontbrekende schakel tussen enerzijds procesanalyse en herontwerp en anderzijds data-analyse. Uitgangspunt voor process mining zijn de zogenaamde event logs. Events worden beschreven door de casus waartoe ze behoren (bijvoorbeeld een patiënt), de naam van de activiteit, een tijdstempel, en een willekeurig aantal aanvullende attributen (bijvoorbeeld de naam van de arts of leeftijd van de patiënt). Dankzij slimme analysetechnieken die zoeken naar veelvuldig voorkomende patronen is het mogelijk automatisch procesmodellen af te leiden ('discovery'). Deze modellen geven inzicht in wat er nu echt gebeurt binnen een proces of organisatie. Vaak wijken de gevonden procesmodellen sterk af van de normatieve procesbeschrijvingen die vaak uitgaan van een ideale situatie die weinig met de werkelijkheid te maken heeft. Om afwijkingen tussen een dergelijke geïdealiseerde procesbeschrijving en de daadwerkelijke gang van zaken in kaart te brengen worden er zogenaamde 'conformance checking'-technieken ingezet. Deze kunnen laten zien wat de mate van compliance is (bijvoorbeeld: '80 procent van de events

verlopen volgens plan') en waar in het proces de grootste afwijkingen te vinden zijn (bijvoorbeeld: 'deze controleactiviteit wordt veelvuldig overgeslagen'). Indien procesmodel en event log redelijk overeenstemmen is het ook mogelijk het procesmodel te verrijken met additionele informatie uit de log. Het is bijvoorbeeld mogelijk in het procesmodel te laten zien waar de knelpunten zitten (dit op basis van tijdstempels in de log).

### Olifantenpaden

Om uit te leggen wat process mining is, wordt vaak de metafoor van een olifantenpaadje gebruikt. Een olifantenpaad is een niet officieel fiets- of wandelpad dat mensen creëren als ze een stukje van de officiële route afsnijden. Doordat ze hun sporen in het gras of plantsoen achterlaten wordt het werkelijke gedrag zichtbaar. Op eenzelfde wijze kunnen process-miningtechnieken op basis van digitale sporen zichtbaar maken wat mensen, systemen en organisaties daadwerkelijk doen. Het process-miningspectrum is echter veel breder

dan het ontdekken van olifantenpaadjes. Het is bijvoorbeeld mogelijk om te ontdekken waarom mensen afwijken van het officiële pad, en waarom delen van het proces langer duren. Olifantenpaden kunnen gebruikt worden om informatiesystemen te verbeteren en mensen op een betere manier aan te sturen. Het doel van process mining is niet om processen in een keurslijf te duwen. Vaak wijken mensen met goede redenen af van het reguliere proces. Hiervan kunnen erg veel organisaties leren. Ook kunnen process-mining-technieken gebruikt worden om te voorspellen wat er gaat gebeuren en om aanbevelingen te doen over de juiste beslissing op basis van historische data. In een ziekenhuis kunnen bijvoorbeeld capaciteitsproblemen voorspeld worden en kan er een alarm afgaan op het moment dat onacceptabele afwijkingen plaatsvinden. Gereedschappen als ProM (zie kader 'Software') ondersteunen dit hele spectrum en kunnen dus daadwerkelijk (Big) data omzetten in betere processen en systemen. <<

## TRADITIONELE DATAMINING-TECHNIKEN SCHIETEN TEKORT

### TOEPASSINGEN

Process-miningtechnieken zijn al succesvol toegepast in honderden organisaties, uiteenlopend van gemeenten en ziekenhuizen tot de leveranciers van hightech-systemen (ASML, Philips en andere). Tijdens het Process Mining Camp op 28 mei ([fluxicon.com/camp/2013](http://fluxicon.com/camp/2013)) kwamen ongeveer honderd gebruikers van process mining bij elkaar. Tijn van der Heijden van Deloitte liet zien hoe process mining de Rabobank in staat stelde om de betalingsprocessen te stroomlijnen. Process mining wordt hier ingezet om de kosten te reduceren en kwaliteit van de afhandeling te verbeteren. Mitchell Cunningham presenteerde een op process mining gebaseerde analyse van de claimprocessen binnen Suncorp, Australië's grootste verzekeraar. Het gaat hierbij om 600.000 claims per jaar die samen meer dan 50 miljoen events genereren. Youri Soons van de Auditdienst Rijk (de interne auditdienst van de rijksoverheid) liet zien hoe afwijkingen en onregelmatigheden in subsidieverleningstrajecten (bijvoorbeeld WBSO) snel opgespoord kunnen worden. Philipp Horn van Volkswagen liet zien hoe alle inkoopprocessen van de verschillende merken van Volkswagen AG (Volkswagen, Audi, Seat en Skoda) vergeleken en geanalyseerd kunnen worden met process mining. Deze voorbeelden laten zien dat het gebruik van process mining snel toeneemt en dat in veel organisaties al voldoende event logs beschikbaar zijn voor gedegen analyses.