



De ontbrekende schakel tussen BI en BPM

De basisconcepten van process mining

De exponentiële groei van gedigitaliseerde data heeft steeds meer invloed op de inrichting en aansturing van bedrijfsprocessen.

Dankzij process mining is het mogelijk om op een gestructureerde manier eventdata te analyseren en te gebruiken voor échte procesverbeteringen.

Wil van der Aalst en Frank van Geffen

Big data staat volop in het nieuws. Alleen organisaties die slim gebruik weten te maken van deze tsunami aan gegevens zullen overleven. Organisaties weten vaak dat waardevolle antwoorden verborgen zitten in hun informatiesystemen, maar traditionele analysegereedschappen zijn niet in staat deze boven water te halen. Process-miningtechnieken zijn wel in staat automatisch procesmodellen te genereren waardoor eventdata ('big or small') kunnen worden omgezet in nieuwe inzichten en betrouwbare antwoorden. Bovendien zijn er steeds betere gereedschappen beschikbaar om dit snel en efficiënt te doen. Dit themanummer van *Informatie* laat zien dat process-miningtechnieken en -tools volwassen genoeg zijn om op grote schaal toegepast te kunnen worden. In dit introductieartikel behandelen we de basisconcepten van process mining. Dit als voorbereiding op de andere bijdragen waar concrete process-miningcasestudies behandeld worden.

Klassieke Business Process Management (BPM)-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 business-intelligencetechnieken tekort vanwege de afwezigheid van een duidelijke procesoriëntatie. Eenvoudige dashboards en beslisbomen kijken vaak naar geaggregeerde data of een enkel punt in het totale proces. Zoals **figuur 1** laat zien is process mining de ontbrekende schakel tussen enerzijds procesanalyse en herontwerp en anderzijds data-analyse.

De meest tot de verbeelding sprekende vorm van process mining is 'process discovery': het automatisch afleiden van procesmodellen (bijvoorbeeld een Petri-net of BPMN-schema) uit eventdata. Dit is echter slechts één van de vele mogelijke analyses. Process mining omvat het hele spectrum van analysetechnieken waar events direct gekoppeld worden aan activiteiten in een processchema. Ook het ontdekken van knelpunten en afwijkingen, het kwantificeren van compliance op basis van echte data, en het voorspellen van doorlooptijden behoren tot het spectrum van process mining.

De aandacht voor data science en big data illustreert de relevantie van het onderwerp. We verwachten dat de behoefte aan data scientists in de komende jaren enorm zal toenemen. Zie bijvoorbeeld de groeiende vraag uit het bedrijfsleven en het ontstaan van onderzoekscentra zoals het Data Science Centrum Eindhoven (DSC/e,

Samenvatting

Process mining is de ontbrekende schakel tussen enerzijds procesanalyse en herontwerp en anderzijds data-analyse. Process-miningtechnieken zijn in staat automatisch procesmodellen te genereren waardoor eventdata ('big or small') kunnen worden omgezet in nieuwe inzichten en betrouwbare antwoorden. Process mining maakt het mogelijk om de effectiviteit, efficiëntie en risicobeheersing van operationele processen feitelijk te analyseren.

www.tue.nl/dsce/). Process mining zal een essentieel gereedschap zijn in de gereedheidskist van de data scientist: uiteindelijk gaat het toch om betere of efficiëntere processen en niet om de data zelf.

Startpunt: eventlogs

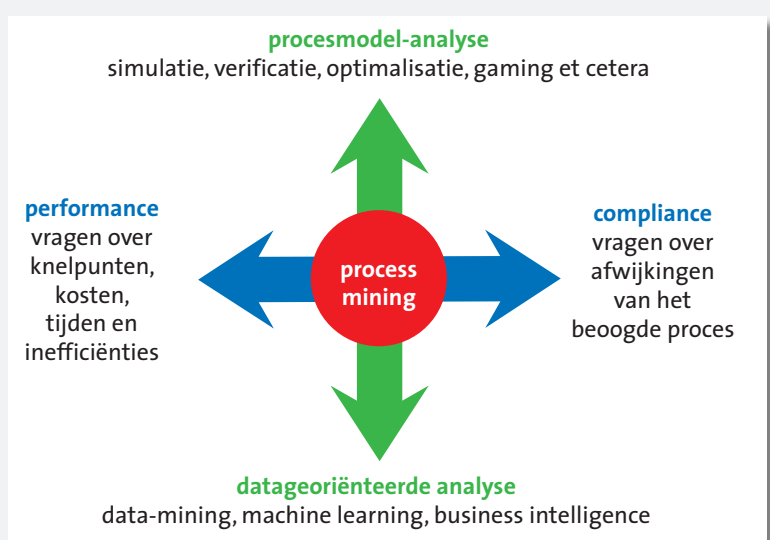
Uitgangspunt voor process mining zijn de zogenaamde eventlogs. Events worden beschreven door de case waartoe ze behoren (bijvoorbeeld een ordernummer), de naam van de bijbehorende activiteit, een tijdstempel, en een willekeurig aantal aanvullende attributen (bijvoorbeeld de naam van de klant of de waarde van de order).

Figuur 2 laat een klein fragment van een eventlog zien (zes events). Elke regel komt overeen met een event. De eerste drie kolommen bevatten de minimale informatie die nodig is voor process mining. Elke event dient naar een case te verwijzen. In dit geval is het patiëntnummer de case-identificatie. Ook dient elke event een naam te hebben (de activiteit) en een tijdstempel. De drie overige kolommen in figuur 2 laten additionele kenmerken van een event zien. Deze kunnen gebruikt worden om bijvoorbeeld de oorzaak van een ontdekt knelpunt of veelvuldige afwijking te bepalen.

Informatie zoals te zien is in figuur 2 treffen we in elk informatiesysteem. Denk bijvoorbeeld aan SAP waar dit soort informatie in diverse tabellen te vinden is. Ook zelfgebouwde of domeinspecifieke systemen bevatten gewoonlijk een schat aan event-informatie. De uitdaging is niet de syntactische conversie naar eventlogs maar het lokaliseren van de data behorende bij een specifiek proces.

Verschillende soorten

Dankzij slimme analysetechnieken die zoeken naar veelvuldig voorkomende patronen is het mogelijk automatisch procesmodellen af te



Figuur 1. Process mining als verbindende schakel tussen model- en data-gebaseerde analyses en performance- en compliance-gerelateerde vragen

patiëntnummer	activiteit	tijdstip	medewerker	locatie	kosten
992345	consult	20-3-2014@10.30	Elle Jansen	Afdeling C	150
995643	consult	20-3-2014@11.00	Elle Jansen	Afdeling C	150
992345	bloedtest	20-3-2014@11.05	Piet Thijsen	Afdeling B	50
992345	röntgenfoto	20-3-2014@11.15	Laura van Aa	Afdeling X	300
995643	bloedtest	20-3-2014@11.25	Piet Thijsen	Afdeling B	50
998765	consult	20-3-2014@11.30	Elle Jansen	Afdeling C	150
...

Figuur 2. Fragment van een eenvoudige eventlog met de case-identificatie (patiëntnummer), de uitgevoerde activiteit en het tijdstip van de event. De overige kolommen zijn optioneel

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



Rabobank

Steeds meer Nederlands organisaties maken de stap om process mining daadwerkelijk toe te passen. Een mooi voorbeeld hiervan is de Rabobank. Rabobank is al sinds 2009 aan het experimenteren met de toepassing van process mining. Twee innovatie-cycli later wordt process mining nu succesvol toegepast in combinatie met: (1) klantinzicht-onderzoeken, (2) Lean Six Sigma en continue verbeteren, (3) operational audits en (4) user experience en systeemgebruik (CRM / ERP). Voor Rabobank is dit een paradigma shift in de wijze waarop besluitvorming- en veranderprocessen zich voltrekken. Ze worden gebaseerd op feiten, volledige data, werkelijke data en snelheid. Dit stelt de Rabobank in staat flexibel en snel te reageren op veranderingen.

gang van zaken in kaart te brengen worden er zogenaamde 'conformance checking'-technieken ingezet. Deze 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 eventlog 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) en waarom cases een bepaald pad volgen (bijvoorbeeld: "Patiënten behandeld door deze arts hebben vaker complicaties").

We onderscheiden twee soorten eventdata: gebeurtenissen van cases die al afgerond zijn en dus niet meer beïnvloed kunnen worden (historische data) en gebeurtenissen van cases die nog niet afgerond zijn (actuele data). Voor de laatste categorie van cases kunnen we met process-miningstechnieken voorspellingen genereren (bijvoorbeeld de verwachte resterende tijdsduur) en aanbevelingen geven (bijvoorbeeld de beste passende medewerker). Dit is mogelijk dankzij de procesmodellen afgeleid van historische data. Normatieve procesmodellen kunnen gebruikt worden om waarschuwingen te genereren op het moment dat er afwijkingen plaatsvinden. Op dit moment richten de meeste process-mining-gereedschappen zich nog op het automatisch afleiden van modellen ('process discovery') en het analyseren van knelpunten en afwijkingen. De verwachting is echter dat process mining ook steeds vaker operationeel (real time) toegepast gaat worden.

Software: let op!

Het opensource-gereedschap ProM ([cessmining.org\) is de de facto standaard op het gebied van process mining in de wetenschappelijke wereld. Onderzoeksgroepen in Nederland, Duitsland, Frankrijk, China, Spanje, Italië, Portugal, Korea, 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 opensource-gereedschap zijn er meer dan zeshonderd plug-ins beschikbaar die het gehele process-miningspectrum afdekken. ProM is echter vooral bedoeld voor process-miningexperts en minder geschikt voor beginnende eindgebruikers.](http://www.pro-</p>
</div>
<div data-bbox=)

Commerciële gereedschappen zoals Disco (Fluxicon), Perceptive Process Mining (eerder Futura Reflect), Celonis, ARIS Process Performance Manager (Software AG), QPR Process Analyzer, Interstage Process Discovery (Fujitsu), Discovery Analyst (StereoLOGIC), en XMAnalyzer (XMPro) zijn eenvoudiger te bedienen en 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. Gereedschappen als Disco kunnen direct CSV-files inlezen en omzetten in het standaard formaat XES (eXtensible Event Stream). XES is toegesneden op process mining en maakt het delen van eventdata eenvoudig.

Vele grotere softwareleveranciers claimen onder de vlag van 'Business Process Intelligence' of 'Business Activity Monitoring' soortgelijke analyses aan te bieden. Dit is echter vaak niet het geval omdat de gebruiker nog steeds handmatig de processen moet modelleren of er geen directe koppeling is tussen de events en activiteiten in het procesmodel. Daarom is het zaak altijd een demonstratie op basis van nieuwe, zelf aangeleverde, data te vragen.

Procesbeheersing en optimalisatie

Process mining maakt het mogelijk om effectiviteit, efficiëntie en risicobeheersing van operationele processen feitelijk te analyseren. Om als organisatie te starten met process mining, is een 'sense of urgency' nodig, gericht op het verkrijgen van inzicht in de echte werking van de operationele processen. De noodzaak ontstaat doorgaans door wijzigingen in de omgeving van het bedrijf, zoals:

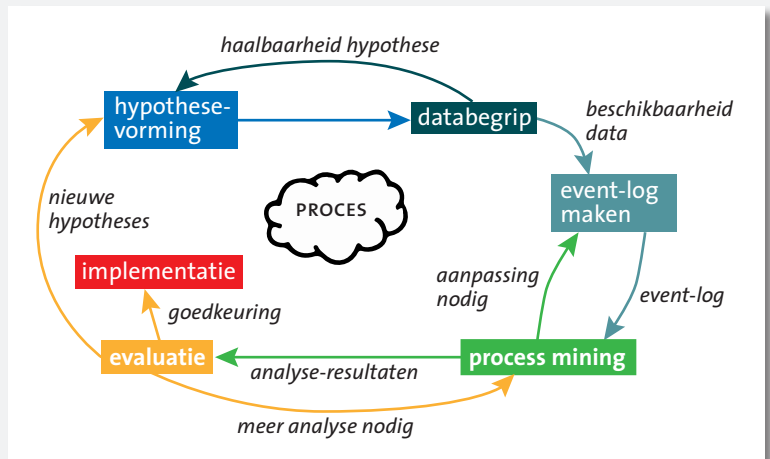
- maatschappelijke ontwikkelingen;
- marktontwikkelingen;
- technologische ontwikkelingen;
- ontwikkelingen in wet- en regelgeving.

De huidige economische crisis creëert een gewijzigde vraag en aangepaste wet- en regelgeving. Maatschappelijk ontwikkelt zich een sterke drang naar transparantie, toezicht en handhaving. Deze ontwikkelingen worden gevoed door een explosie van data, analyse- en visualisatietechnieken waarmee snel, eenvoudig en transparant inzicht wordt verkregen. Voor organisaties betekent dit een noodzaak tot het verbeteren van hun aanpassingsvermogen en objectieve, transparante verantwoording. En dus een focus op operationele processen. Deze focus kan langs twee assen worden bekeken: (1) procesoptimalisatie (verbetering en of vernieuwing) en (2) procesbeheersing (compliance).

De eerste as is hoofdzakelijk een verantwoordelijkheid van de business. Procesoptimalisatie wordt primair gedreven door wensen van de klant in balans te brengen met de wensen van de organisatie. Process mining draagt bij aan het feitelijk inzichtelijk maken van bijvoorbeeld ongewenst lange doorlooptijden, teveel handelingen, verspilling en rework. De tweede as is het aandachtsgebied van de 'internal control' en auditfunctie. Vanuit een bedrijfsbrede risicoanalyse worden kritische processen benoemd. Process mining stelt de auditor in staat een feitelijk en objectief beeld te verkrijgen over de werking van de kritische processen en de mate van beheersing.

Morgen starten

Figuur 3 geeft de globale lifecycle van process-miningprojecten weer (Van der Heijden 2012). Het biedt een mooie kapstok om te beschrijven hoe je als organisatie kunt starten met process mining. Dit model komt overeen met de levenscyclus uiteengezet in het *Process Mining Manifest* (2012). Voor de succesvolle



Figuur 3. Lifecycle van process-miningprojecten

»Uiteindelijk gaat het om betere of efficiëntere processen en niet om de data zelf«

aanpak van process mining is het echter zaak gewoon te beginnen. De benodigde data zijn overal te vinden en het is vooral een kwestie van doen. Dit wordt geïllustreerd in de verdere casestudies in dit themanummer.

Wil van der Aalst

Van der Aalst is universiteitshoogleraar aan de Technische Universiteit Eindhoven (TU/e). Daarnaast heeft hij deeltijdaanstellingen bij *Queensland University of Technology (QUT)* in Australië en de *National Research University Higher School of Economics* in Moskou. Hij is ook trekker van de IEEE Task Force on Process Mining. E-mail: w.m.p.v.d.aalst@tue.nl

Frank van Geffen

Van Geffen is process innovator bij de Rabobank en voorzitter SIG Process Mining bij de Ngi-NGN. E-mail: f.geffen@rn.rabobank.nl

Literatuur

- W.M.P. van der Aalst (2011). *Process Mining: Discovery, Conformance and Enhancement of Business Processes*. Heidelberg: Springer.
- IEEE Task Force on Process Mining (2012). *Process Mining Manifesto*. In *Lecture Notes in Business Information Processing* 99, pages 169-194. Springer-Verlag.
- W.M.P. van der Aalst, A. Adriansyah, en B. van Dongen (2012). *Replaying History on Process Models for Conformance Checking and Performance Analysis*. *WIREs Data Mining and Knowledge Discovery*, 2(2):182-192.
- T.H.C. van der Heijden (2012). *Process Mining Project Methodology: Developing a General Approach to Apply Process Mining in Practice*, Technische Universiteit Eindhoven, alexandria.tue.nl/extra2/afstversl/tm/Van_der_Heijden_2012.pdf.
- M. Haasnoot Bezverhaya en F. van Geffen (2013). *Process Mining en Internal Auditing*. *AUDIT magazine, Magazine voor internal en operational auditors*, 12(4).