

ARTIKELN | 16 DECEMBER 2020

GRIP OP DE ALGORITMES IN DE ORGANISATIE

Auteur: Emil van Werven MSc EMITA - Gerke Roorda MSc EMITA - Pim Smulders MSc RE CISA

Beeld: Erik Mclean

Leestijd: 8 min



Het gebruik van algoritmes neemt hand over hand toe, als vervanging van of als aanvulling op bestaande applicaties of rekenmodellen. Het gemak en de snelheid van algoritmes zorgen ervoor dat ze niet meer weg te denken zijn. Hoe krijg je grip op de algoritmes in de organisatie?

Vrijwel elke organisatie zal algoritmes vroeg of laat gaan inzetten. De inzet brengt ook onderhoud, beheer en beheersing van risico's met zich mee. Op dit moment staat het beheersen van de risico's ten aanzien van algoritmes nog in de kinderschoenen. Regelmatig komen bedrijven en overheidsorganisaties in het nieuws wanneer er iets is misgegaan met de inzet van een algoritme.

Beheersen van algoritmes

Hoe krijg je grip op de algoritmes in de organisatie? Waar moet je op letten als internal audit? Hoe voorkom je dat je zaken over het hoofd ziet? In hoeverre wijkt het beheersen van algoritmes af van de beheersing van andere IT-risico's? De eerste stap is natuurlijk de inventarisatie van waar en op welke wijze algoritmes gebruikt worden in de organisatie. Deze inventarisatie vormt vervolgens de basis van een risico-analyse. Iedereen voelt op zijn klompen aan dat de risico's van een algoritme ten behoeve van een spamfilter anders zijn dan een algoritme die tweets analyseert met het oog op

geautomatiseerd beleggen. Hoe voer je gefundeerd een dergelijke risicoanalyse uit en bepaal je welke algoritmes wel en welke niet in scope dienen te zijn voor internal audit? En hoe bepaal je weloverwogen welke beheersingsmaatregelen je tenminste zou moeten implementeren?

In dit artikel gaan we eerst dieper in op de karakteristieken van een algoritme. Wat maakt een algoritme nu anders dan de beheersing van andere IT-systemen of applicaties? Wat zijn typische risico's waarmee je bij de inzet van algoritmes rekening dient te houden? Daarna beschrijven we een door ons ontwikkeld hulpmiddel: een generiek normenkader dat houvast biedt om gefundeerd een risicoanalyse uit te voeren en te komen tot een stelsel van maatregelen voor de beheersing van deze risico's.

Karakteristieken van een algoritme

Een algoritme wordt vaak ingezet om een (gecompliceerde) beslissing geautomatiseerd te nemen of daar geautomatiseerd ondersteuning bij te bieden. Ten behoeve van het nemen van die beslissing worden grote hoeveelheden data verwerkt om te komen tot een kansinschatting van het rekenmodel. Daarbij heeft een algoritme ook de kenmerkende factor dat het veranderlijk van aard kan zijn: de berekeningen in het rekenmodel, de meegenomen variabelen en weging daarvan kunnen veranderen gedurende de inzet van een algoritme, omdat het zichzelf kan wijzigen om maximaal resultaat te behalen.

Dat klinkt zowel aantrekkelijk als gevaarlijk, want aan de ene kant hoeft de ontwikkelaar niet meer over alle factoren na te denken en kan de rekenkracht van machines efficiënter worden ingezet. Dat maakt dat algoritmes wendbaar kunnen zijn in een veranderende omgeving. Aan de andere kant is het ook goed mogelijk dat een algoritme afdrijft en over de tijd steeds minder goede beslissingen gaat nemen. Er zijn situaties denkbaar waarbij het op hol slaan van zo'n algoritme niet wordt opgemerkt door degene die dit inzet. Het is belangrijk om daarbij nog op te merken dat een algoritme zo goed presteert als de grenzen die het heeft meegekregen door zijn 'schepper'.

Vroeg of laat krijgt elke organisatie te maken met de inzet van algoritmes. Hoe pak je de audit ten aanzien van algoritmes aan?

Gevolgen voor risico's

Een aantal van de karakteristieken van algoritmes leiden ertoe dat de risico's rondom een algoritme wezenlijk anders kunnen zijn dan de gebruikelijke risico's die we verwachten bij bijvoorbeeld 'klassieke' rekenmodellen. Een belangrijk risico is inherent aan de kansinschatting wat de output is van het algoritme. Bij het inschatten van een kans is het natuurlijk ook mogelijk om foutief te 'gokken'. Binnen de ontwikkeling van algoritmes wordt dan gesproken over false positives en false negatives: voorspellingen die onterecht tot een positieve of negatieve uitslag of beslissing hebben geleid.

We spraken eerder al over de grote hoeveelheden data die worden ingezet bij het voorspellen van volgende waarden, de uitkomsten van een algoritme. Bij gebruik van datasets komt het begrip bias al snel ter sprake; een ongewenste afwijking in de dataset die ervoor zorgt dat de data niet representatief is voor de populatie waar deze op geprojecteerd wordt. Bias raakt ook al snel ethiek, met name op het moment dat het verschil tussen man/vrouw, ras of nationaliteit een rol speelt in de populatieprojectie. Zo kan het gevaarlijk zijn om op basis van een dataset met meer mannen dan vrouwen als datapunten conclusies te trekken; bij een evenwichtige dataset zou die conclusie immers anders kunnen zijn. Het is praktisch onmogelijk om alle risico's daarin te definiëren. Echter, het is wel goed mogelijk om bias en verschillende randvoorwaarden voor bias als risico te definiëren.

Dan is er nog een typisch risico rondom de inzet van algoritmes waar lastig grip op te krijgen is. Door de toepassing van, soms verschillende, complexe technieken en het gebruiken van de rekenkracht van machines om te komen tot een adequate weging van variabelen in een model, kan het bijzonder lastig zijn om uit te leggen wat nu heeft geleid tot een

(positieve of negatieve) voorspelling. Bij het voorspellen van het favoriete gerecht van een willekeurig persoon op basis van een aantal variabelen is de impact van dat risico beperkt, maar op het moment dat een bank niet kan uitleggen waarom zij een klant een hypotheek weigert, wordt dat risico cruciaal in de acceptatie van de technologie en de beheersing ervan.

Overige risico's

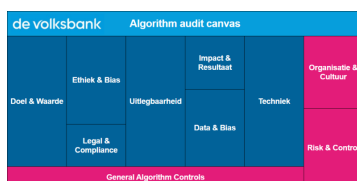
Ten opzichte van business rules en rekenmodellen zijn er ook risico's die weliswaar zeer relevant zijn voor algoritmes, maar niet per se afwijken van wat gewoon is in de inschatting van risico's rondom meer traditionele technologieën. Rondom bias is al benoemd dat ethische aspecten een rol kunnen spelen. Dat geldt ook voor veel meer toepassingen van grote datasets. Specifiek van algoritmes weten we niet altijd of daar sprake van is of niet. Maar ook in een veel bredere context is ethiek een belangrijk aspect in de toepassing van technieken, waar vragen bij gesteld kunnen worden als: willen wij wel besluiten nemen op basis van deze en deze variabelen? Of: wat is het doel van de toepassing en past dat bij wat wij als organisatie ethisch verantwoord vinden?

Bij die tweede vraag komt dus ook het doel en de waarde van het algoritme naar voren. Het is goed om bij (nieuwe) technologische toepassingen bewust stil te staan bij wat die toepassing oplevert en of dat gewenst is. Als dat niet diepgaand genoeg wordt nagegaan voorafgaand aan de ontwikkeling, ontstaan daar serieuze risico's. Tegelijkertijd heeft iedere toegepaste techniek ook haar eigen risico's, waar rekening mee gehouden zal moeten worden.

Naast risico's die specifiek relevant worden bij het toepassen van algoritmes, bestaan er ook generieke risico's, zoals rondom de beheersing van risico's, het opstellen en voeren van beleid en voldoen aan wet- en regelgeving en interne procedures en richtlijnen. Daarnaast zijn de klassieke onderwerpen zoals changemanagement, cyber security en logische en fysieke toegangsbeveiliging (alle vallend onder de general IT controls) ook van toepassing op IT-omgevingen waarin algoritmes zich bevinden. Deze risicogebieden zijn dus ook van toepassing op algoritmes, net als de specifieke risico's, en zullen dus meegenomen moeten worden in een risico-analyse daarop.

Algemeen normenkader: het algorithm audit canvas

Hoe zijn al deze risico's en aandachtsvelden te vangen in een generiek normenkader? Om dit in kaart te brengen hebben we ons laten inspireren door het business model canvas. Met behulp van een zevental nationale en internationale richtlijnen (onder andere Oxford-Munich code of conduct, impact assessment AI ECP en ethics guidelines for trustworthy AI) is een generiek algorithm audit canvas ontwikkeld, bestaande uit tien verschillende aandachtsvelden (zie *figuur 1*). De groene vlakken in het figuur hebben direct betrekking op het proces waarin het algoritme functioneert. De rode vlakken zijn randvoorwaardelijk. Per aandachtsveld zijn beheersingsdoelstellingen met bijbehorende risico-aandachtsgebieden geformuleerd. Hierna per aandachtsgebied in het canvas een korte beschrijving van de inhoud.



Figuur 1. Algorithm audit canvas

Doel & waarde – Waarvoor wordt het algoritme gebruikt en wat is de toegevoegde waarde van het gebruik hiervan voor zowel de organisatie als de omgeving waar het algoritme invloed op heeft.

Ethiek & bias – Betreft welke (mogelijke) ethische dilemma's het gebruik van het algoritme kan oproepen binnen de organisatie, de stakeholders en de maatschappij. Door bepaalde besluitvorming over te laten aan een algoritme kunnen

er fundamentele waarden in het geding komen. Hierbij is het van belang dat de organisatie rekening houdt met deze waarden om de maatschappij niet te schaden.

Uitlegbaarheid – In de beheersing rondom een algoritme dienen voldoende aanknopingspunten te zijn om in begrijpelijke taal de inzet, het doel en de uitkomst van het algoritme te kunnen uitleggen. Deze aanknopingspunten betreffen bijvoorbeeld de mate waarin stakeholders betrokken zijn bij de ontwikkeling en inzet van het algoritme.

Legal & compliance – Betreft de wet- en regelgeving waarmee een organisatie rekening dient te houden bij het gebruik van een algoritme. Tevens gaat het in op de maatregelen die organisaties hebben getroffen om te borgen dat het algoritme voldoet aan de eisen die zijn gesteld vanuit de wet- en regelgeving.

Impact & resultaat – De impact van een algoritme kan zeer breed zijn. Het is dan ook van belang dat wordt vastgesteld dat het algoritme accurate resultaten oplevert.

Data & bias – Gaat in op de data die het algoritme gebruikt om tot een oordeel te komen. Hierbij is het van belang dat de data van voldoende kwaliteit zijn om het algoritme goed te kunnen laten werken en het de juiste output te laten genereren.

Techniek – Betreft de onderliggende technologie waarop het algoritme werkt; de code. Er zijn verschillende typen algoritmes te onderscheiden die elke weer op hun eigen manier werken. Daarom is het van belang de onderliggende techniek mee te nemen in het beheersen van het algoritme, aangezien iedere techniek een eigen aanpak heeft.

Risk & control – Belicht de risico's die het gebruik van een algoritme met zich meebrengt en op welke wijze deze door middel van een PDCA-cyclus worden beheerst. Belangrijk is dat de organisatie maatregelen treft om risico's die het gebruik van een algoritme met zich meebrengt te mitigeren, om zo het gebruik ervan op organisatieniveau te beheersen.

Organisatie & cultuur – Gaat in op de relatie tussen de inrichting van de organisatie, de aanwezige cultuur en de inzet en het gebruik van algoritmes. Hierbij valt zowel te denken aan beleidsmatige stukken die zijn opgesteld waarin de visie beschreven is en het benoemen en beleggen van processen en verantwoordelijkheden. Daarnaast heeft dit betrekking op de wijze waarop binnen de organisatie vorm wordt gegeven aan de cultuur als bedding voor het gebruik en de inzet van een algoritme binnen of door de organisatie.

General algorithm controls – Belichten de randvoorwaardelijke beheersmaatregelen (bijvoorbeeld toegang tot de data en het changemanagement van het algoritme) die nodig zijn om te borgen dat een algoritme werkt zoals deze behoort te werken. Indien hierin beperkingen worden geconstateerd, zou dit impact kunnen hebben op de werking van het algoritme. Daarom is het van belang dat voor deze randvoorwaardelijke aandachtsgebieden voldoende maatregelen zijn getroffen om te borgen dat deze goed zijn ingericht.

Diepgang canvas en raakvlakken met andere professies

Tijdens een evaluatie van het canvas is de diepgang ervan uitvoerig besproken. In onze benadering van het concept algoritme richten wij ons niet specifiek op de techniek en code van het algoritme, maar vooral op het hele proces en bijbehorende beheersing. Wanneer de auditor van oordeel is dat de technische aspecten van het algoritme met meer diepgang onderzocht moeten worden, dan dienen deze werkzaamheden uitgevoerd te worden door een daartoe bekwame data scientist, aangezien een IT-auditor in de regel niet over deze expertise beschikt. Een IT-auditor heeft de verantwoordelijkheid om een afweging te maken tot welke grens hij voldoende expertise heeft om werkzaamheden zelf uit te voeren en vanaf welk punt een data scientist moet worden ingeschakeld.

Conclusie

AUDIT

MAGAZINE

We begonnen dit artikel met de vraag hoe je gefundeerd een risicoanalyse uitvoert om te bepalen welke algoritmes in scope dienen te zijn bij internal audit en hoe je vervolgens weloverwogen bepaalt welke beheersingsmaatregelen tenminste geïmplementeerd zouden moeten zijn. Vanwege de veelzijdigheid van algoritmes en omdat een algoritme wezenlijk anders is dan gewone rekenmodellen of applicaties, komt er een breed scala aan risico's in beeld. Wij pretenderen niet alle risico's geïnventariseerd te hebben. Wel hebben wij vanuit een eigen analyse, aangevuld met zeven nationale en internationale richtlijnen, een generiek canvas ontwikkeld. Dit canvas is een hulpmiddel om relevante risico's in kaart te brengen en weloverwogen keuzen te maken. Het zou ons niet verbazen als in de komende tijd nog risico's naar voren komen die wij niet onderkend hebben. Dat is inherent aan nieuwe technologie. Wij stellen het canvas graag beschikbaar voor wie er gebruik van wil maken om het in te zetten in de organisatie en/of verder door te ontwikkelen. Want vroeg of laat krijgt elke organisatie te maken met de inzet van algoritmes. En daarmee komt ook de vraag op elke internal auditafdeling af: hoe pak je de audit ten aanzien van algoritmes aan?

Over

Emil van Werven, Gerke Roorda en Pim Smulders zijn allen actief bij Baker Tilly IT Advisory als senior consultant, waar zij zich vanuit het externe beroep bezighouden met IT-auditopdrachten. Zij deden in het kader van hun afstudeerproject voor de Executive Master of IT Auditing aan TIAS Tilburg in het voorjaar 2020 onderzoek naar de risico's rondom uitlegbaarheid van algoritmes.

Tags: Risicomanagement, RPA

Bron url: <https://auditmagazine.nl/artikelen/grip-op-de-algoritmes-in-de-organisatie/>