

<Whitepaper>

De data-gedreven accountant aan zet

Door: Renee Bonestroo



Inhoud

Inhoud	2
Inleiding	3
De moderne accountant	4
In 5 stappen naar meer regie	6
Inventariseer:	6
Waar liggen data kansen:	7
Kies een geschikte use-case:	7
Start, experimenteer en leer:	8
Evalueer en start met een volgende use-case:	8
Use-case: Datagedreven adviseren	9
Inleiding use-case	9
Voorbeeld use-case:	10
Huidige situatie	10
Vorbereiding	11
Technische inrichting	11
Eindelijk, aan de slag met data!	12
Data gedreven adviseren	13
De datagedreven organisatie	15
Wat is een Data Mesh?	15
Nieuwe termen	17
Waarom zou je kiezen voor een Data Mesh?	18
Hoe start je dan vervolgens met een Data Mesh?	18
Hoe richt ik een dataplatform in?	20
Waar begin je dan?	20
Selecteren van de bouwblokken	21
Conclusie	25

Inleiding

Data wordt steeds belangrijker, ook in de wereld van de accountant. De technologische ontwikkelingen gaan hard, de hoeveelheid data neemt toe. Handmatig controles uitvoeren wordt steeds lastiger. Ook ontstaat er in toenemende mate vraag om data aan te leveren aan organisaties als SIRA of t.b.v. AQI's. In hoeverre is de accountant klaar om mee te gaan met deze veranderende omstandigheden en is de accountant klaar om gebruik te maken van alle ontwikkelingen op gebied van technologie en data.

Data wordt nu regelmatig gestald bij een derde partij waardoor data niet direct beschikbaar is en het toevoegen van nieuwe informatiebronnen stroperig en kostbaar is. Ook zien we dat dataverwerking niet altijd op een efficiënte manier gebeurt en dat data kansen niet voldoende benut worden. Is de moderne accountant klaar om regie te pakken met data?

Wij vanuit yellow arrow vinden van wel, maar daar is wel verandering voor nodig. Graag vertellen wij meer over deze verandering en hoe je als moderne accountant de data-regie in handen kan pakken.

De moderne accountant

Met de komst van de cloud wordt technologie toegankelijker voor het grote publiek. Zo wordt dataopslag steeds goedkoper en wordt rekenkracht gescheiden van dataopslag waardoor je rekenkracht op een efficiëntere manier kunt inzetten. Tegelijkertijd is er in de accountancy branche een belangrijke ontwikkeling gaande en dat is dat steeds meer data uitwisseling gestandaardiseerd wordt. Denk hierbij aan uitwisseling van jaarrekeningen, audit files en bankbestanden waar afspraken over gemaakt worden zodat data zoveel mogelijk op een uniforme manier uitgewisseld kan worden. Dit schept de ideale condities om informatiebronnen zelf, zo veel mogelijk uniform en geautomatiseerd, te gaan ontsluiten en een eigen analyse omgeving op te zetten.

Waarom zou je energie steken in het opbouwen van een eigen analyse omgeving? De accountant heeft vaak al een hoge werkdruk en heeft misschien niet de vaardigheden van een dataspecialist. De reden om hier wel tijd en energie in te steken is omdat het op de lange termijn veel oplevert. We noemen hier een aantal voordelen:

Pak de regie

- Als accountant zit je zelf aan de knoppen – Jij bepaalt wat je nodig hebt, welke data relevant is en welke niet. Wil je berekeningen op een andere manier doen, wil je data vanuit een andere perspectief bekijken, jij bepaalt en jij weet hoe je jouw klant het beste kan bedienen.
- Afhankelijkheid van derde partijen wordt kleiner – De data wordt verzameld in een eigen omgeving. Jij bepaalt welke data wordt toegevoegd en welke niet. Je hoeft niet weken te wachten op een derde partij die dat vervolgens weer met verschillende andere partijen en klanten moet afstemmen.
- Altijd toegang tot jouw data – Data hoeft niet op aanvraag beschikbaar worden gesteld. Zodra de data in jouw analyse omgeving ingeladen is, is het ten alle tijden direct beschikbaar. Voorkom dat je 'gegijzeld' wordt door softwareleveranciers. Door data zelf op te slaan houd je de beschikking over de data ook wanneer software contracten worden beëindigd.

Strategie

- Data uit verschillende bronnen combineren – In de praktijk is data vaak gefragmenteerd opgeslagen. Een accountant krijgt te maken met veel verschillende applicaties en verschillende bestandsuitwisselingen. Al deze data verzamelen is tijdrovend en kan ook zeer foutgevoelig zijn. Wat als je deze data geautomatiseerd in een eigen analyse platform kan ontvangen?

Efficiency

- Repeterend werk wordt geautomatiseerd zodat je kan focussen op de echte waarde – Wie zit er niet uren in Excel data te prepareren, draaitabellen te verversen, macro's op nieuw te

draaien etc. Hoe fijn zou het zijn als je het repeterende werk kan automatiseren, zodat je meer tijd over houdt voor andere belangrijke werkzaamheden.

- Kwaliteit controles verbeteren – Door controles te automatiseren kunnen twee voordelen worden behaald: Er kunnen grotere hoeveelheden data gecontroleerd worden en daarnaast kan de kwaliteit van controles verbeterd worden. Doordat er minder handmatig gecontroleerd wordt gaat de foutmarge omlaag en ook kunnen nieuwe technieken zoals AI en machine learning controles aanzienlijk verbeteren. Denk bijvoorbeeld aan fraude detectie op basis van slimme algoritmes.
- Met een juiste kostenbeheersingsstrategie zijn de consumptie kosten in een cloud oplossing laag te houden – We horen vaak dat organisaties bang zijn dat kosten in een cloud omgeving heel hard kunnen oplopen. Daar hebben ze niet direct ongelijk in, maar met de juiste maatregelen en inrichting kunnen kosten wel degelijk laag gehouden worden.
- Mogelijkheid om veel meer data te verwerken - De toename van data zorgt er voor dat traditionele oplossingen niet meer werken. Een schaalbare, cloud gebaseerde analyse omgeving maakt het mogelijk om analyses op grote schaal te doen

Gevolg

De moderne accountant pakt zelf de regie, heeft toegang tot data wanneer nodig en heeft de vaardigheden om data-analyses uit te voeren. Data-gedreven werken is wat ons betreft de toekomst. Er zijn veel discussies gaande bij bedrijven of ze de stap moeten maken naar (meer) data-gedreven werken. Is dit wel nodig? Daar hebben wij een simpel antwoord op; Ja dat is nodig. De wereld draait om data, en er komt alleen maar meer data bij. We zitten midden in een transitie waarin elk bedrijf nu of in de nabije toekomst data-gedreven zal gaan werken. Medewerkers die daar niet in mee willen, maar ook organisaties die daar niet in mee willen zullen steeds minder relevant worden. Het moment is nu, ook in de accountancy branche, om daarop in te haken.

Yellow arrow

Natuurlijk zijn wij ook een beetje van 'wc-eend'. Wij zien dat er de nodige stappen te zetten zijn om de traditionele accountant te transformeren naar een moderne data-gedreven accountant. Wij zetten graag onze kennis en kunde in om daarbij te helpen. In het volgende hoofdstuk vertellen we meer over onze aanpak.

In 5 stappen naar meer regie

Oké, je ziet de voordelen van een eigen analyse platform. Maar waar begin je? En wat komt er allemaal bij kijken? Dit zijn belangrijke vragen! Het is niet verstandig om te beginnen zonder enige voorbereiding. Je zult niet de eerste zijn die vol goede moed een data project start om vervolgens volledig te verdwalen in de wereld van data.

Een goede voorbereiding en een goede aanpak zijn daarom belangrijk. Hieronder een stappenplan wat jullie de handvatten geeft om een data-initiatief in goede banen te leiden:

- 1) Inventariseer, waar sta je nu?
- 2) Verken data kansen
- 3) Kies een use-case om te starten
- 4) Start, experimenteer en leer
- 5) Evalueer en start met een volgende use-case

Inventariseer:

Om te beginnen is het goed om vast te stellen waar je nu staat. Wat heb je wel- en wat heb je niet in huis. Hieronder drie belangrijke onderdelen om van tevoren vast te stellen:

- Wat voor kennis en kunde is er aanwezig binnen jouw organisatie?

Om (meer) data gedreven te worden als organisatie heb je data bedreven medewerkers nodig. Een project valt of staat met de adoptie van jouw medewerkers. Wie in jouw organisatie zijn de data bedreven medewerkers? Wie zijn erin staat om data te analyseren met oplossingen als PowerBI of een andere BI-tool. Daarnaast belangrijk om vast te stellen wie er technisch kennis hebben om een dataoplossing te ontwerpen en te beheren? Is deze kennis in huis of heb je een partner die hierbij kan helpen?

- Welke data tools zijn er binnen jullie organisatie beschikbaar?

Werk je nu voornamelijk in Excel of heb je Business Intelligence tools als bijvoorbeeld Power BI of QlickView waarmee je analyses kan doen? Wordt de analyse uitbesteedt aan een derde partij die jouw data beheerd en analyses toestuurt? Hoe makkelijk heb je nu toegang tot de benodigde data en is dat voldoende?

- Welke databronnen zijn er?

De digitale transitie is al even gaande. Ook accountantskantoren hebben veel geïnvesteerd in systemen voor klantgegevens, klant documentatie en financiële gegevens. Daarnaast ontvangt een accountant data van klanten, handmatig dan wel geautomatiseerd. Wat wordt er allemaal uitgewisseld en in hoeverre zijn deze uitwisselingen van data gestandaardiseerd? Kunnen we dit automatiseren?

Waar liggen data kansen:

Het verkennen van data kansen is onderdeel van de inventarisatie fase, maar vanwege het belang van dit onderdeel beschouwen we dit als aparte stap in het proces. Het verkennen van kansen is een proces wat je niet vanuit een IT-bril moet bekijken. Dit is ook niet iets wat je als afdeling alleen moet doen. Kansen verkennen is een creatief proces wat je vooral samen en afdeling overstijgend moet aanvliegen. Plan 1 of meerdere brainstormsessies en kies zorgvuldig de mensen die daar een zinvolle bijdrage aan kunnen leveren. Maak er ook geen Poolse landdag van door de halve organisatie uit te nodigen. Ongeveer 6 tot 10 medewerkers uit verschillende domeinen is aan te raden. Nodig medewerkers uit met uitgebreide domeinkennis en affiniteit met data-analyse. Geef de medewerkers echt de ruimte om data kansen te onderzoeken en zorg dat jouw medewerkers enthousiast worden over data. Alles valt of staat met draagvlak binnen jouw organisatie.

Wanneer je de juiste mensen hebt samengebracht probeer dan de kandidaat use-cases te identificeren en op waarde te schatten. Op waarde schatten is belangrijk om de juiste keuzes te kunnen maken en prioriteit aan te brengen. Het opstellen van use-cases is een techniek uit de software engineering met als doel het gedrag van een systeem, in dit geval een data product, te beschrijven vanuit het perspectief van de gebruiker. Een use-case laat zien wat een dataproduct moet toevoegen voor zijn gebruikers om doelen te bereiken. Een use-case kan ook helpen om de scope, de functionaliteiten en de requirements van een dataproduct te definiëren. Doe dit in eenvoudige en begrijpelijke taal, zonder technische details. In een later moment wordt de vertaling naar de techniek gemaakt. Wanneer je dit zonder technische details uitwerkt laat je je ook niet beperken door eventuele onmogelijkheden van technologie.

Mocht je het als organisatie lastig vinden om kansen te identificeren of wil je graag geholpen worden in dit proces? Vanuit yellow arrow hebben wij een workshop ontwikkeld om samen met jullie eigen domeinspecialisten deze kansen te verkennen, op waarde te schatten en te prioriteren. Wil je meer informatie over de workshops neem dan contact met ons op! [Contact | yellow arrow](#)

Kies een geschikte use-case:

Kies een use-case waarvan je goed kan beoordelen dat het waarde oplevert en waarvan jouw medewerkers of klanten ook de voordelen van zien. Iets kan financiële waarde opleveren maar hoeft niet direct voordelen te hebben voor jouw medewerkers of klanten. Wanneer je een use-case selecteert welke voor jouw medewerkers of klanten direct waarde toevoegt creëer je draagvlak voor data gedreven werken.

Hoe bepaal je dan deze waarde en waarom is dit zo belangrijk? De succesvolle data gedreven organisaties onderscheidt zich in drie zaken:

1. Zet data succesvol in binnen de dagelijks besluitvorming,
2. Kan de waarde creatie van en door data meten
3. Senior management is volledig gecommitteerd om de datareis succesvol te maken

Kortom om succesvol te zijn in jouw datastrategie is het belangrijk dat jouw organisatie in staat is de waarde te meten. Dat begint bij het inschatten wat een use-case aan waarde kan creëren. Kijk hierin eerst naar het laaghangend fruit. Hoeveel uren besparen jouw medewerkers door het implementeren van deze use-case? Het aantal uur kan je vertalen naar een kostprijs en je hebt een eerste indicatie. Andere voordelen zijn misschien niet direct in geld uit te drukken. Bijvoorbeeld wanneer je een betere klant tevredenheid verwacht. Hier zul je creatiever mee om moeten gaan. Stel

een baseline op en meet, bijvoorbeeld daar klanttevredenheid onderzoek of direct klant contact, of de dienstverlening in de loop van de tijd verbeterd.

Selecteer nu aan de hand van deze waarde-inschatting de juiste use-case.

Start, experimenteer en leer:

Wanneer je kan onderbouwen dat een use-case echte waarde toevoegt ga dan vooral aan de slag. Controleer samen met jouw domeinspecialisten en technische mensen nog wel even of de juiste technische randvoorwaarden aanwezig zijn om te kunnen starten. Onderdeel van deze fase is het opzetten van een technische architectuur, waarbij een ervaren data engineer of een data architect de vertaling maakt van de use-case naar een technische oplossing. Hierbij kijkt de architect of engineer naar verschillende zaken zoals het kostenplaatje (initieel en recurring), functionaliteit van de technische bouwblokken, verwachte leercurve van de medewerkers die er mee aan de slag moeten, schaalbaarheid en verwachte onderhoudsinspanning.

Wanneer de technische oplossing staat kan een data bedreven medewerker eventueel samen met een data engineer data gaan verkennen, verzamelen en verreiken om de use-case te realiseren. Dit is een leerproces en zal absoluut met vallen en opstaan gaan. Niet elk initiatief is direct een schot in de roos. Leer van jouw datareis, deel je bevindingen met collega's. Begin opnieuw wanneer je vastloopt. Het is echt beter om overnieuw te starten, dan doorbouwen op een slecht fundament. Het is ook niet erg om kennis en kunde in de markt te zoeken. De juiste skills extern aan te haken is in eerste instantie een extra kostenpost, maar deze kosten zul je terugverdienen doordat je op de lange termijn minder leergeld kwijt zult zijn.

Evalueer en start met een volgende use-case:

Wanneer de use-case opgeleverd wordt evalueer met het betrokken team de resultaten. Ook hier geldt: Fouten maken is niet erg zolang je er maar van leert. Neem ook vooral mee wat goed ging naar de realisatie van een volgende use-case. Door een dataplatform use-case voor use-case op te bouwen kan je kort cyclisch werken en snel bijsturen indien dit nodig mocht zijn. Dit is de agile gedachte waar het draait om leveren van waarde, snel bijsturen, leren en verbeteren terwijl je bezig bent.

Use-case: Datagedreven adviseren

We hebben het gehad over de voordelen van een eigen analyse omgeving gehad en de voordelen zijn wat ons betreft niet te ontkennen. Om daar meer gevoel bij te krijgen schetsen we een use-case waarmee de toegevoegde waarde van een eigen analyse omgeving tastbaarder wordt.

Inleiding use-case

De gekozen use-case beschrijft een voorbeeld situatie gebaseerd op het fictieve bedrijf Geel & van der Pijl Accountants. Dit voorbeeld is gebaseerd op onze ervaringen met data gerelateerde projecten in de wereld van de accountant. Wij zijn technologie/data specialisten en geen accountants, maar hebben in onze samenwerking met accountants wel verschillende inzichten opgedaan die wij graag met jullie delen.



Een kerntaak van een accountant is het verzamelen van gegevens voor het controleren van financiële cijfers. Deze cijfers worden verzameld om de juistheid, betrouwbaarheid en integriteit van de boekhouding te controleren. Ook heeft een accountant een verantwoordelijkheid bij het detecteren van fraude. Naast deze belangrijke taken heeft een accountant ook een adviserende rol. Dat wil zeggen dat een accountant op basis van de geanalyseerde data een klant kan adviseren over fiscale zaken, processenverbeteringen, financiële problemen of symptoombestrijding van een ongezonde organisatie.

Het is goed om te beseffen dat de wereld van de accountant en het administratiekantoor veranderd. De meeste betalingen gaan nu volledig automatisch. Ook het facturatieproces kan bij veel bedrijven in hoge mate worden geautomatiseerd door boekhoudsoftware. Dit zorgt ervoor dat jaarrekeningen automatisch samengesteld kunnen worden en dat de boekhoudsoftware ook al veel zorgen omtrent juistheid en integriteit kan wegnemen. Hiermee wordt de rol van de accountant anders en komt de nadruk veel meer te liggen op de adviesrol van een accountant. Het is wel goed om hier duidelijk onderscheid te maken tussen de samenstel- en de controlepraktijk. In de controlepraktijk gelden andere regels met betrekking tot het advies geven en is er altijd een spanningsveld tussen de mate van adviseren en het onafhankelijkheidsbeginsel.

Maar zoals gezegd neemt boekhoudsoftware steeds meer taken van de accountant weg, maar dezelfde accountant zal nog steeds bankbestanden, jaarrekeningen en audit bestanden moeten verwerken vanuit de controlepraktijk. Het verwerken van deze data kan ook voor een groot deel geautomatiseerd worden, helemaal vanwege de ontwikkeling dat data uitwisseling van bankbestanden, jaarrekeningen en audit bestanden gestandaardiseerd en geüniformeerd wordt. Voorbeeld hiervan zijn het cash management (CAMT^{*1}) bestanden die opgevraagd kunnen worden bij een bank. Deze standaard is vastgelegd in de ISO 20022 standaard en alle banken in Nederland conformeren zich aan deze standaard.

Nog meer automatisering dus. Blijft er dan nog wel wat over voor de accountant? Worden

*1 : CAMT staat voor Common Architecture for Messages for Transfers en wordt gebruikt om elektronische afschriften uit te wisselen tussen banken en klanten.
<https://www.iso20022.org/iso-20022-message-definitions>

accountants dan niet met de tijd overbodig? Wij denken absoluut van niet, maar de moderne

accountant richt zich veel meer op adviseren op basis van data-analyse en om dat te kunnen zul je als moderne accountant de data regie stevig in handen moeten pakken.

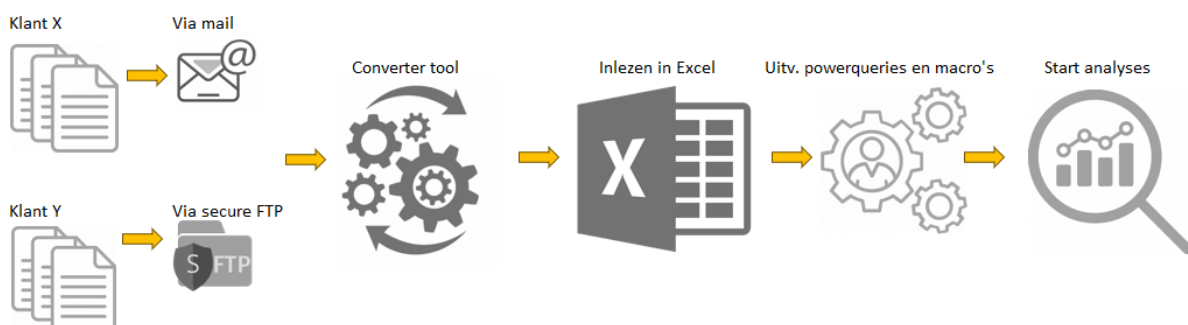
Voorbeeld use-case:

Geel & van der Pijl Accountants wil graag zelf de regie over data. Het idee is ontstaan om zelf data in een eigen analyse omgeving te gaan opslaan. Er zijn nu verschillende medewerkers die data opvragen of toegestuurd krijgen van banken of klanten. De aanleveringen zijn veelal volgens marktstandaarden en daarmee goed te automatiseren. Het betreft hier data welke verzameld wordt vanuit de controlepraktijk. Door de data in een eigen analyse omgeving op te slaan worden de volgende voordelen verwacht:

- Per controle de benodigde preparatie tijd met 3 manuren (gemiddeld) inkorten. Hierdoor kan er veel sneller met de data-analyses gestart worden.
- Significant minder fouten doordat het proces geautomatiseerd wordt. Hierdoor worden controle en correctie werkzaamheden met gemiddeld 4 uur per analyse teruggebracht. Controles en correcties zijn binnen deze context gericht op het proces om de aangeleverde data te prepareren en in zijn volledigheid in Excel te krijgen. Het gaat hier nog niet om de inhoudelijke controle van de aanlevering.
- De tool om aangeleverde bestanden naar Excel om te zetten is niet meer nodig en hiermee kunnen licentie kosten worden uitgespaard.
- Klantgerichtere analyses en klantgerichte advisering. Hier zit de echte winst en dit is dan ook het belangrijkste voordeel. De tijd die vrijkomt door deze automatisering wil Geel & van der Pijl Accountants inzetten om voor haar klanten betere analyses te ontwikkelen en om klanten gericht te kunnen adviseren.
- 1 versie van de waarheid. Door regie te voeren over de data en door business logica centraal vast te leggen gebruikt iedere dataconsument dezelfde definities en ontstaat er 1 versie van de waarheid. Ook geen verschillende Excel kopieën meer.
- Betere beveiliging van vertrouwelijke informatie. Duidelijke rollen en rechten, data maskering van privacygevoelige data en encryptie van data moet de kans op ongeautoriseerde toegang verkleinen. Dit is in Excel een stuk lastiger te waarborgen.

Huidige situatie

Binnen het huidige proces bestaat er een Excel template waar data ingeladen kan worden om analyses te doen. Voordat de bestanden in Excel verwerkt kunnen worden, wordt er eerst een externe tool ingezet om de data te converteren van hiërarchische bestandsstructuren als XML en JSON naar bruikbare platte lijsten welke makkelijk in Excel ingeladen kunnen worden. In het Excel template zitten verschillende powerqueries en macro's om de data te prepareren om vervolgens de



data in draaitabellen beschikbaar te stellen waarna gestart kan worden met de inhoudelijke analyses. In de praktijk gaat het verversen niet altijd goed. De powerqueries of macro's gaan nog wel eens wat stuk. Ook het verversen van macro's en formules in Excel gaat niet altijd direct goed. Daarnaast worden bestanden gedeeld onder collega's die elk in hun eigen versie aan de slag gaan. Hierdoor ontstaan verschillende waarheden. Vanuit de bril van de accountant is dit een ongewenste situatie omdat je data op een eenduidige manier wilt verwerken. Meerdere versies van de waarheid kan afbraak doen aan de betrouwbaarheid van de accountant en controles bemoeilijken.

Er wordt ook vaak de vraag gesteld of data van verschillende aanleveringen van een klant gecombineerd kunnen worden zodat er bijvoorbeeld trendanalyses gedaan kunnen worden. Met deze trendanalyses kan er beter ingespeeld worden op de wensen van specifieke klanten. Bijvoorbeeld wanneer een klant terugkerende cashflow problemen ondervindt. Door een trendanalyse te doen op hun data kunnen verbanden ontdekt worden waardoor er een beter beeld ontstaat wanneer en waardoor deze cashflow problemen ontstaan. Hierdoor kan de accountant een onderbouwd advies geven over de te treffen maatregelen en voorspellen wanneer er weer problemen kunnen gaan ontstaan als er geen adequate mitigerende maatregelen worden getroffen.

Sommige klanten leveren data aan via een beveiligde ftp-omgeving, andere klanten leveren bestanden aan via de mail. Er worden vragen gesteld of het wel verstandig is dat sommige klanten data aanleveren via de mail. Weliswaar worden bestanden met wachtwoorden beveiligd, maar deze manier van beveiligen is door een kwaadwillende relatief eenvoudig te omzeilen. Is een geautomatiseerde en versleutelde verbinding niet beter?

Vorbereiding

Alvorens er een use-case uitgewerkt wordt is de eerste stap in ons stappenplan de inventarisatie van de huidige situatie. Welke kennis is in huis, welke tooling hebben we tot onze beschikking en welke bronssystemen hebben we nodig. Uit deze inventarisatie bij Geel & van der Pijl Accountants is gebleken dat zij over een aantal behoorlijk bedreven data analisten beschikken, maar dat technische kennis om een analyse omgeving op te zetten ontbreekt. Naast Excel zijn er geen analyse/BI tools beschikbaar maar is daar wel behoefte aan.

Deze inventarisatie is nodig om een goed plan te kunnen maken en om commitment vanuit het management af te dwingen. Enerzijds heb je funding nodig voor bijvoorbeeld nieuwe BI tooling en externe inhuur en anderzijds om draagvlak te creëren om als organisatie meer data-gedreven te gaan werken.

Na een goede presentatie aan het MT, waar de voordelen en verwachte kosten goed onderbouwd konden worden, krijgen de initiatiefnemers de ruimte en de funding om aan de slag te gaan met dit initiatief. Zoals aangegeven in onze vorige blog en onderdeel van ons stappenplan is het belangrijk dat je de verwachte opbrengst kan inschatten en onderbouwen. Dit is ook weer onderdeel van het creëren van draagvlak, maar is zeker niet altijd even makkelijk. Ook een goede inschatting maken van de verwachte kosten, met name wanneer er sprake is van externe inhuur, kan lastig blijken. Wij vanuit Yellow Arrow hebben ervaring in het onderbouwen van use-cases en zetten graag onze kennis hiervoor in.

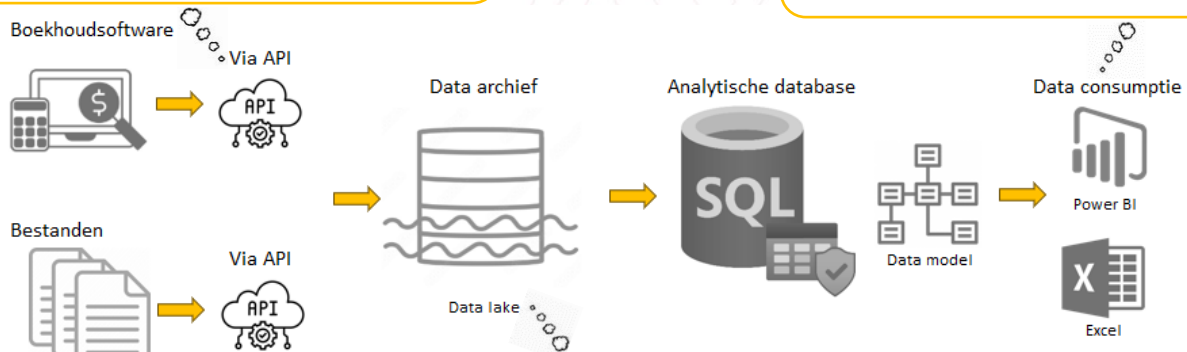
Technische inrichting

Omdat Geel & van der Pijl Accountants niet de juiste technische kennis in huis heeft wordt ervoor gekozen om een technologie/data partner te zoeken om de technische inrichting te ontwerpen en te realiseren. Samen met deze technologie/data partner wordt er een data-analyse platform in een cloud omgeving ontworpen. In dit voorbeeld gaan we uit van een oplossing binnen Microsoft Azure

om de use-case te realiseren. Hoe je een ontwerp van een modern analyse platform maakt binnen een cloud-omgeving naar keuze komt aan bod in het vierde deel van deze blog-reeks. Om deze use-case te realiseren is het volgende ontwerp bedacht. Dit is ontstaan uit een intensieve samenwerking tussen de kennishouders, data analisten en de technologie/data partner:

1) Data wordt geautomatiseerd opgehaald middels API's. Een **Application Programming Interface**, afgekort tot API, is een verzameling definities op basis waarvan een computerprogramma kan communiceren met een ander programma of onderdeel.

4) In een BI tool als Power BI kan de verrijkte data opgehaald en gevisualiseerd worden. De business logica op het platform is voor iedereen hetzelfde en ligt eenduidig vast. 1 versie van de waarheid.



2) De data wordt opgeslagen in een data archief zoals een data lake.

3) Door een analytische database wordt de data verwerkt, geïntegreerd en verrijkt. Tevens wordt hier het historisch beeld gereproduceerd op basis van het archief. Het verrijkte data model kan doormiddel van een BI tool geconsumeerd worden.

Eindelijk, aan de slag met data!

We hebben een mooi initiatief bedacht en uitgewerkt in een use-case, de juiste mensen zijn aangehaakt, we hebben commitment van het management en we hebben een modern analyse platform ingericht in Microsoft Azure. Nu kunnen we de data daadwerkelijk gaan verzamelen. Wanneer je zelf niet de kennis hebt om de informatie te verzamelen binnen een analyse platform zul je in eerste instantie moeten leunen op de dataspecialisten van jouw partner. Dat is niet erg, maar probeer zelf wel de kennis op te bouwen door enerzijds medewerkers op training te sturen of door dataspecialisten als data engineers aan te trekken.

Het verwerken, integreren en verrijken van data is een onderwerp waar we meerdere hoofdstukken aan kunnen wijden, maar voor nu gaan we er gemakshalve van uit dat de data op een juiste manier verzameld wordt zodat de data geconsumeerd kan worden door een BI-tool als Power BI. In het volgende hoofdstuk willen we vooral stil staan bij de voordelen die de realisatie van deze use-case mogelijk maken.

Data gedreven adviseren

Hoe maakt het automatiseren van deze data stromen het mogelijk om beter en data gedreven te gaan adviseren voor Geel & van der Pijl Accountants.

- **Tijdwinst:** Door het automatiseren van de oplossing houden de accountants meer tijd over. Dit kunnen ze inzetten om meer gericht advies te geven en hiermee extra waarde voor hun klanten te genereren.
- **Historie:** Door het opbouwen van historie kan je, naast de reguliere controles en bijbehorende advisering, een totaalbeeld vormen en worden ook andere vormen van advies mogelijk zoals trendanalyses.
 - 1) We noemden al de cashflow analyse waar op basis van trendanalyse cashflow problemen voorspelt kunnen worden. Om een goede trendanalyse te kunnen maken heb je data van een langere periode nodig dan alleen het lopende boekjaar. Deze analyse zou je ook voor andere indicatoren kunnen opzetten zoals bijvoorbeeld solvabiliteit, rentabiliteit of winst/verlies.
 - 2) **Fraude detectie.** Door transacties over langere periodes te monitoren aan de hand van slimme algoritmes kunnen patronen in de data ontdekt worden. Patronen die je niet ziet als je de data zelf analyseert in bijvoorbeeld Excel. Ook als je alleen kijkt naar het lopende boekjaar heb je mogelijk niet voldoende data om daadwerkelijk die patronen te ontdekken om fraude te kunnen opsporen.
 - 3) **Procesmining.** De interimcontrole wordt vaak ingezet als voorbereiding op de eindejaar controle. Een belangrijk onderdeel van de interimcontrole is de procesanalyse. Hoe komt de data tot stand en hoe betrouwbaar is de data. Met technieken als proces mining kunnen procesanalyses worden geautomatiseerd. Met slimme algoritmes kunnen afwijkingen in grote hoeveelheden data worden gedetecteerd. Hiermee kunnen datakwaliteitsproblemen en hiaten in het proces aangetoond worden.
- **Eén versie van de waarheid:** Voor een accountant in de controlepraktijk draait alles om de waarheid en het constateren dat jouw klanten een correcte en juiste administratie voeren. Maar zoals een timmerman die vakwerk aflevert bij klanten, maar thuis geen zin meer heeft om de boel netjes af te timmeren omdat hij de hele dag toch al aan het timmeren is, zo moet je als accountant ook je eigen data huishouding niet verwaarlozen. Een situatie waarbij meerdere personen met kopieën van Excel documenten werken moet je in deze tijd echt niet meer willen. Dit zorgt voor fouten en verschillen. Door deze verschillende kopieën ontstaan meerdere versies van de waarheid. Door business logica eenduidig en centraal vast te leggen in een analyse platform houdt je regie.
- Door de komst van het analyse platform wordt het veel makkelijker om de beveiliging in te richten: Je kan autorisaties geautomatiseerd toekennen op basis van functies. Je kan data maskering toepassen op gevoelige data en je kan data encrypten om de kans op ongeoorloofde toegang te voorkomen. Daarnaast kan je inrichten dat toegang tot de data alleen mogelijk is vanaf vertrouwde devices of vertrouwde netwerken.
- **Benchmarking:** Op basis van geanonimiseerde data is het mogelijk om benchmarks uit te voeren. Hoe verhouden de cashflow, winst/verlies of investeringen van jouw organisatie zich

met andere branchegenoten. Neem jij veel risico, of ben je juist terughoudend ten opzichte van jouw concurrenten. Hoe je deze data presenteert moet je goed over nadenken en helemaal de mate van advies welke je verbindt aan deze cijfers. Hier heb je weer te maken met het onafhankelijkheidsbeginsel van de controlepraktijk en dat mag niet ter discussie staan.

- Risicoanalyse: Door de analyse van de huidige situatie op basis van daadwerkelijke problemen en de daarbij ontdekte patronen, kunnen we risico's identificeren die voortkomen uit deze patronen.

Het zelf verzamelen van data, het opbouwen van een data archief en het inrichten van een omgeving waar de data-analyse uitgevoerd kan worden geeft je de mogelijkheid om een efficiëntie slag door te voeren. Daarnaast geeft dit extra gereedschap om jouw klanten beter te bedienen. Hoe je dit gereedschap straks gebruikt ligt aan jullie specifieke situatie, visie en aanpassingsvermogen. Daarbij zul je ook altijd goed moeten kijken naar actuele wet- en regelgeving. Ons advies is ga vooral met een soortgelijke use-case aan de slag! Laat je adviseren waar nodig en weet dat het datateam van yellow arrow klaar staat om hierbij te helpen.

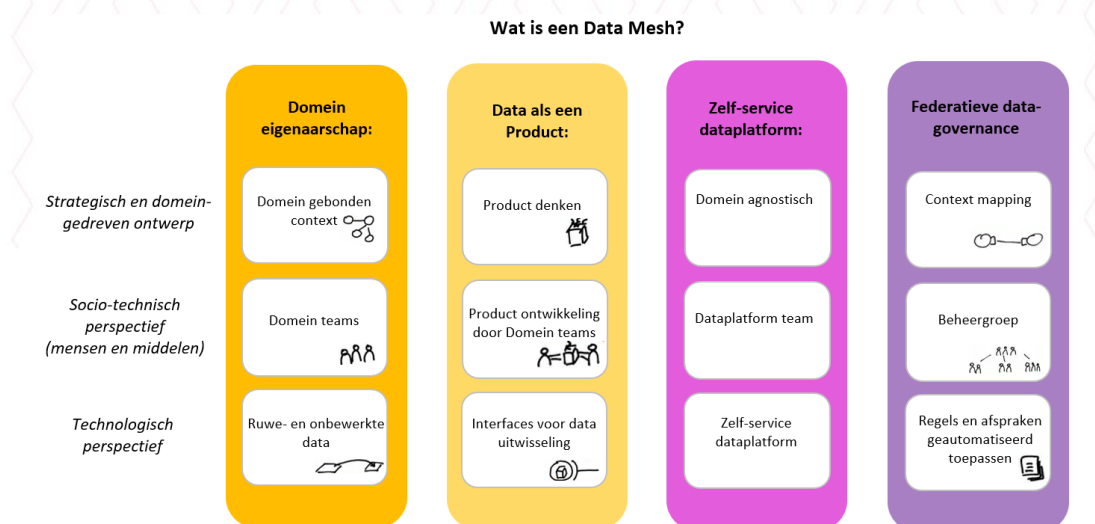
De datagedreven organisatie

We hebben in deze whitepaper stil gestaan bij waarom datagedreven werken zo belangrijk is, ook voor accountants. We hebben ingezoomd op de voordelen van datagedreven adviseren aan de hand van een use-case bij een fictief accountantskantoor (natuurlijk wel gebaseerd op onze ervaringen). En zoals we in ons 'vijf stappenplan naar meer regie' aangaven is dat het verstandig is om een dataplatform use-case voor use-case op te bouwen. Je kan dus klein beginnen en het platform stap voor stap opbouwen. Naarmate het dataplatform volwassener wordt zul je ook goed moeten nadenken over organisatorische aspecten als data governance. Met data governance worden vragen beantwoordt als: Wie is de eigenaar van de data? Wie kan ik vragen stellen als ik de data niet begrijp? Hoe weet ik naar wat voor data ik eigenlijk kijk, m.a.w. wat is de betekenis (definitie) van de data datapunten waar ik naar kijk? Waar komt mijn data vandaan?

Hoe houdt je regie als adoptie van dataproducten groeit binnen jouw organisatie. Of te wel hoe houdt je controle over jouw data wanneer steeds meer collega's zelf met de data aan de slag gaan. Een handige methodiek die hierbij kan helpen is de Data Mesh. Eerst zullen we uitleggen wat een Data Mesh überhaupt is en dan zullen we inzoomen op waarom deze Data Mesh zo'n interessante methodiek is.

Wat is een Data Mesh?

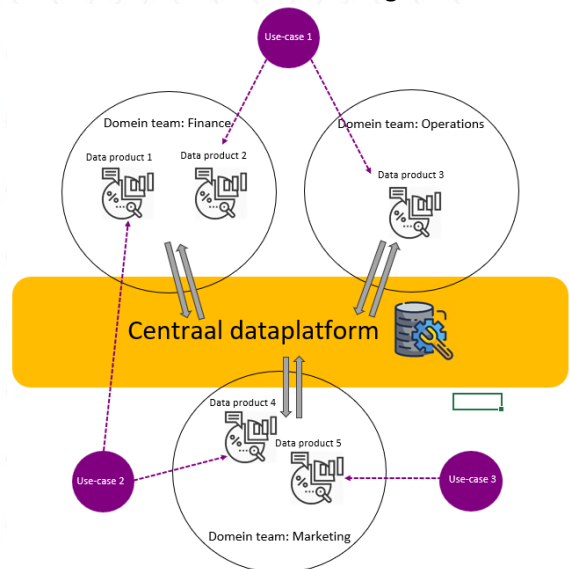
De term Data Mesh werd in 2019 geïntroduceerd door Zhamak Dehghani, Directeur opkomende technologieën bij Thoughtworks. De Data Mesh is een methodiek of beter gezegd een architectuur waarmee (complexe) dataplatformen ingericht kunnen worden om op een zo efficiënte manier, als data gedreven organisatie, data producten uit te nutten en nieuwe data producten te creëren. Enerzijds kan je de Data Mesh zien als een architectuur omdat de Data Mesh sterk bepaald hoe jouw data architectuur eruit komt te zien. Anderzijds kan je het ook als een methodiek zien omdat er ook een sterke focus ligt op organisatorische uitdagingen. Data Mesh richt zich op 4 basis principes:



- **Data eigenaarschap:** In de Data Mesh betekend eigenaarschap dat domein teams verantwoordelijk zijn voor hun eigen data. Dat wil niet zeggen dat iedereen weer in zijn eigen silo, met eigen middelen en technologie data oplossingen gaat bouwen. Nee, het is juist de bedoeling dat de data zoveel mogelijk centraal opgeslagen wordt,

iedereen ook dezelfde werkwijze en standaarden hanteert, maar wel domein specifieke dataproducten ontwikkeld. Als medewerker binnen een bepaald domein ben jij degene die er toch het meeste verstand van heeft? Of je nu medewerker bent in het domein belastingadvies of onderdeel bent van de controlepraktijk.

- Data als een product:** Een tweede belangrijk principe is dat je naar data gaat kijken als een product. Producten die met zorg zijn samengesteld en een hoge kwaliteit hebben. Daarnaast is het belangrijk dat dataproducten gedeeld kunnen worden, zodat medewerkers in andere domeinen data weer kunnen gebruiken. Enerzijds om op te sturen en anderzijds om weer nieuwe dataproducten te ontwikkelen. Zo ontstaat een keten van dataproducten.
- Zelf-service data omgeving:** Binnen een datagedreven organisatie zou iedereen toegang tot relevante data moeten hebben. Zoals we in ons vijf stappenplan al beschreven is het belangrijk dat medewerkers hier wel de juiste competenties voor opbouwen. Het idee dat iedereen zelf met, de voor haar rol of functie, relevante data aan de slag gaat is simpel: Als domein specialist ken jij jouw domein het beste. Een data engineer heeft technische skills, maar heeft zeker niet altijd specifieke domein kennis zoals een belastingadviseur of een register accountant dat wel heeft. Dit betekent dus niet dat we weer terug gaan naar losse data oplossingen. Iedereen werkt wel op het centrale dataplatform en gebruikt dezelfde standaarden en richtlijnen om aan kwaliteitseisen te kunnen voldoen en om te zorgen dat iedereen op een eenduidige manier werkt. Dit heeft als gevolg dat de rol van data engineer langzaam aan het veranderen is en er een nieuwe rol aan het ontstaan is. Dit is de rol van analytics engineer. De data engineer zorgt voor alle koppelingen met bronnen en zorgt dat het fundament van het platform staat. De analytics engineer is een domein specialist met technische bagage. Deze specialisten zorgen dat de data getransformeerd worden naar bruikbare data producten. De eindgebruiker, of te wel de dataconsument kan data producten combineren om vraagstukken mee op te lossen.
- Federatieve data-governance:** De governance van data is cruciaal binnen de Data Mesh. Wil je al deze functionele domeinen goed met elkaar laten samenwerken dan zijn er duidelijke spelregels nodig en ook duidelijke afspraken over kwaliteit: Ook wel de standaarden en de richtlijnen. Naast deze standaarden en richtlijnen moeten er duidelijke afspraken gemaakt worden over het delen van data. Hier komt ook weer een nieuw concept om de hoek, namelijk het data contract. Dit is een afspraak over hoe een dataproduct geconsumeerd kan worden. In het contract liggen aspecten als kwaliteit, periodiciteit van verversing en definities van de data vast. Hierdoor weet de dataconsument precies hoe hij de data kan gebruiken en wat hij er wel of niet van mag verwachten. Door dataproducten te delen voorkom je ook dat business logica vaker ontwikkeld wordt. Je wilt namelijk maar 1 versie van de waarheid. En anders zul je business logica ook op meerdere plekken moeten beheren en onderhouden wat sowieso niet verstandig is. Wanneer er dus een nieuwe data vraag komt kijk je eerst op de menukaart wat

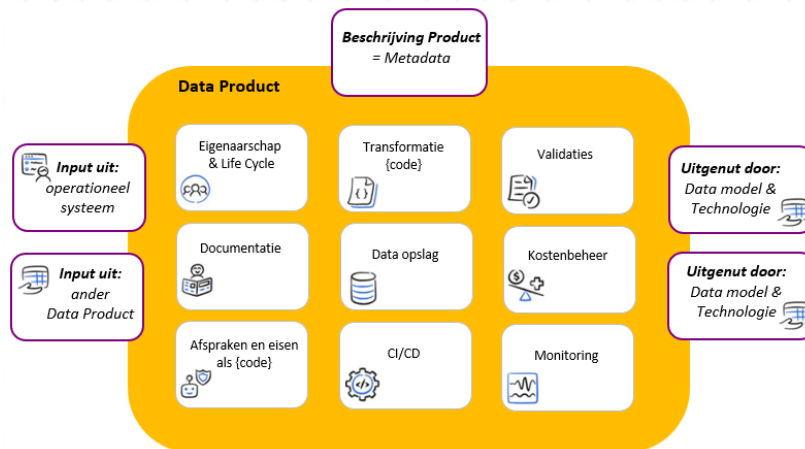


er al is. Daar kan je eventueel eigen ingrediënten aan toevoegen om een nieuw (data-)gerecht te maken. Mocht het een lekker gerecht zijn, dan laat je het gerecht ook weer opnemen op de menukaart.

Nieuwe termen

Binnen de Data Mesh worden een aantal nieuwe termen geïntroduceerd. Voordat we het gaan hebben over waarom je voor een Data Mesh zou moeten kiezen is het goed om even wat verder op deze termen in te zoomen.

Dataproduct: De term dataproduct is niet per se nieuw, maar wel de manier hoe een dataproduct gedefinieerd wordt binnen een Data Mesh. Een Data Product is het logische geheel van onderdelen binnen een dataplatform welke nodig zijn om een use-case te kunnen realiseren. Deze definitie is redelijk abstract, maar betekend concreet dat een dataproduct niet alleen de data zelf betreft, maar ook alle (technische) onderdelen die nodig zijn om de use-case te kunnen realiseren. Hiervoor zijn verschillende bouwblokken nodig zoals data storage (gegevensopslag), code (transformatie en business logica), documentatie, monitoring en logging, security policies en onderdelen voor beheer. Een dataproduct koppelt aan bronsystemen of aan andere dataproducten. Het dataproduct voegt waarde toe middels nog niet eerder ontwikkelde business logica. Deze logica is goed beschreven, beheer en onderhoud is ingericht en de kwaliteit en beschikbaarheid worden continue geborgd. Dat ziet er schematisch zo uit:



Data Contract: Een data contract definieert de voorwaarden waaronder een dataproduct afgenomen kan worden. Denk hierbij aan hoe de data wordt aangeleverd (welk formaat), de kwaliteit van de data en hoe deze kwaliteit gegarandeerd wordt (m.a.w. wat voor controles worden er gedaan), de definities van data punten, eigenaarschap en periodiciteit van verversen. Eventueel kunnen ook de kosten van het dataproduct opgenomen worden in het contract zodat deze doorbelast kunnen worden aan dataconsumenten.



Federatieve data governance: Zoals reeds benoemd is data governance cruciaal in de Data Mesh. Data governance begint met het maken van heldere afspraken over hoe dataproducten en data

contracten worden opgezet. Begin met afspraken maken over hoe je data kan vastleggen en delen binnen jouw organisatie. Bepaal hoe je definities kan vastleggen en bij wie je moet zijn als er inhoudelijke vragen zijn over de data. Er komen ook steeds meer hulpmiddelen die hierbij kunnen helpen. Denk bijvoorbeeld aan Azure PurView, Data Mesh Manager en DBT waar je dataproducten, data contracten en zaken als eigenaarschap en definities kan vastleggen.

Waarom zou je kiezen voor een Data Mesh?

Bij het toelichten van de 4 peilers van de Data Mesh is misschien al duidelijk geworden wat een Data Mesh aan voordelen biedt. De belangrijkste uitdaging die de Data Mesh probeert op te lossen is het knelpunt van een centrale data team. Nu steeds meer bedrijven stevig inzetten op datagedreven werken neemt de vraag naar dataproducten toe. De centrale datateams hebben het hierdoor ontzettend druk. Het duurt steeds langer om data beschikbaar te krijgen omdat de backlog groeit en groeit. Daarnaast er is er ook meer behoefte aan data uit ander soort bronnen waar data niet- of op een andere manier gestructureerd is. Denk hierbij aan sensor en machine data of video en spraak. De hoeveelheid van data neemt hierdoor ook significant toe (hierdoor is de term 'big data' ook ontstaan). Door meer in te zetten op zelf-service kan je het centrale datateam ontlasten.

De dataconsument verwacht ook een kant- en klare dataproduct. De data engineer is al superdruk en dan moet de data engineer ook nog eens in de tijd die hij over heeft het domein van de dataconsument goed proberen te begrijpen om met een dataproduct te komen wat naadloos aansluit op de wensen van de consument. Dit is bijna onmogelijk. Door de domein specialisten in staat te stellen om zelf dataproducten te maken lossen wordt deze tweede uitdaging ook op.

Door de Data Mesh wordt de organisatie in staat gesteld om bij te blijven bij snel veranderende omstandigheden in de markt en technologie. Dit is nodig om de ambitie om (meer) datagedreven te gaan werken waar te kunnen maken. Ook helpt de Data Mesh om iedereen in zijn of haar kracht te zetten. Dus een data engineer in het verzamelen van data en een domein specialist om relevante informatie uit die data te halen.

Hoe start je dan vervolgens met een Data Mesh?

Allereerst is het verstandig jezelf de vraag te stellen of Data Mesh in jullie huidige situatie en met het huidige volwassenheids niveau ook daadwerkelijk iets toevoegt. En dit is dan ook meteen een erg lastige vraag om te beantwoorden, maar vanuit yellow arrow proberen we IT simpeler te maken. Dus dat gaan we hier ook doen aan de hand van de volgende vragen:

1. Hoeveel domeinen binnen jouw bedrijf kan je identificeren die sterk afhankelijk zijn van data of waar data een duidelijke toegevoegde waarde kan hebben.
2. Is er al een centraal datateam en vormt dit datateam een bottleneck bij dataoplossingen? Op een schaal van 1 tot 10, waarbij 1 'nooit' en 10 'altijd' betekend.
3. Heeft data governance prioriteit binnen jouw organisatie? Met andere woorden is er behoefte en inzicht in zaken als eigenaarschap, betekenis- of herkomst van data. Op een schaal van 1 tot 10 waarbij 1 'geen idee wat data governance uberhaupt is' en 10 'ja enorm behoefte aan meer grip op data en datastromen' betekend.
4. Hoe belangrijk is kwalitatief goede data voor het nemen van beslissingen binnen jouw organisatie (van strategisch tot operationeel). Ook weer op basis van een schaal van 1 tot 10

waarbij 1 'beter slechte data dan geen data' en 10 'betrouwbare en correcte data is cruciaal voor onze operatie en de beslissingen die wij nemen' betekend.

Tel de uitkomst van de vier vragen bij elkaar op. Scoor je minder dan 10 punten dan is Data Mesh waarschijnlijk nu nog niet interessant voor jouw organisatie. Scoor je tussen de 10 en 30 punten dan is het goed om over Data Mesh na te denken. Begin met verdiepen in de Data Mesh en waar mogelijk pas al op kleinschalige manier concepten en best-practice toe. Scoor je boven de 30 punten dan is ons advies om serieus met Data Mesh aan de slag te gaan, omdat jouw organisatie sterk kan profiteren van de voordelen van de Data Mesh.

Word je al enthousiast over de Data Mesh en zie je de voordelen binnen jouw organisatie? Dat is mooi! Wacht dan ook niet tot het ideale moment zich voordoet om te starten. De kans is klein dat er überhaupt een ideaal moment komt. Ook hier geldt weer ga er vooral mee aan de slag, start klein (maar denk groot) en kies een geschikt pilotproject. Advies is ook om met 1 domein te beginnen. Je zult eerst een basis moeten neerzetten om later andere domeinen te kunnen laten aanhaken. Deze basis bestaat uit de volgende onderdelen:

- Maak een on-boardingsplan voor de zelf-service medewerkers: Daar worden vragen beantwoord als: Wat betekend zelf-service data binnen jouw organisatie? Welke skills moeten de zelf-service medewerkers hebben? En hoe worden de medewerkers opgeleid?
- Leg de spelregels vast van de Data Mesh: Wat is een dataproduct? Waar moet een dataproduct aan voldoen? Wat zijn de standaarden en richtlijnen om deze dataproducten op een juiste en uniforme manier te ontwikkelen. En hoe borgen we de kwaliteit?
- Hoe gaat data gedeeld worden zowel binnen- als buiten domeinen? Hier komt ook het data contract om de hoek kijken. Beschrijf wat een data contract is en hoe data contracten worden afgesloten.
- Wat betekend deze federatieve data governance voor jouw organisatie? Wat brengt dat voor eisen met zich mee met betrekking tot het voorbereidingsproces van data?
- Wat is de impact op het huidige dataplatform. Moet je jouw data architectuur aanpassen om het zelf-service principe goed te kunnen ondersteunen. En moet er extra tooling komen om bijvoorbeeld data contracten vast te leggen en zijn de juiste middelen en tools aanwezig om deze federatieve data governance te kunnen voeren?

Vanuit yellow arrow hebben wij de kennis en ervaring om te ondersteunen bij het implementeren van een Data Mesh. Wil je hier graag meer over weten laat het ons weten [Contact | yellow arrow](#)

Hoe richt ik een dataplatform in?

Als laatste hoofdstuk binnen dit whitepaper willen we stil staan bij, vanuit het oogpunt van de accountant misschien wel het minst leuke onderdeel, de techniek. We hebben het al over de noodzaak van data gedreven werken gehad. Waar liggen kansen en wat betekent dit voor jouw organisatie. Om echt data-gedreven te gaan werken en die regie in handen te pakken heb je een eigen dataplatform nodig. Daar komt wel een stukje techniek bij kijken. De ‘grote vier’ beschikken over eigen datateams en zetten vol in op hun data competenties, maar dat is niet voor elk accountantskantoor al het geval. Nu is het grote voordeel wel dat met de komst van cloud omgevingen als Microsoft Azure en Amazon Web Services (AWS) de drempel veel lager is geworden om een eigen omgeving op te zetten. Ook de kosten, wel met een degelijk ontwerp en mits je de juiste maatregelen neemt, kunnen flink lager zijn dan traditionele oplossingen. Daarnaast kan je klein beginnen en investeringen ook stapsgewijs nemen waardoor het voor kantoren met een minder grote omvang ineens ook heel interessant wordt om een eigen dataplatform op te zetten. Wat ons betreft is een modern dataplatform in de cloud dan ook de juiste weg vooruit en onze aanpak is ook volledig op een cloud oplossing gericht.

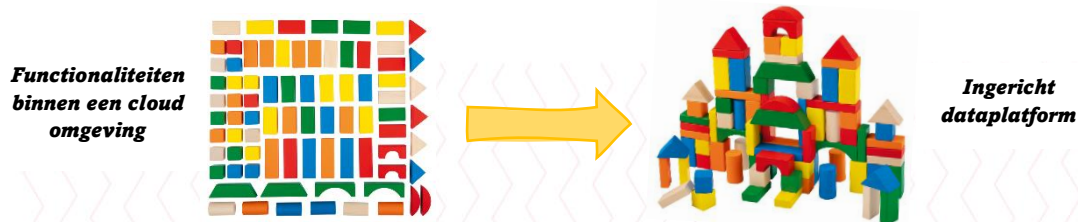
Waar begin je dan?

Laten we weer terugkijken naar het stappenplan uit hoofdstuk 1. Het begint met het verkennen van data kansen vanuit een business perspectief. De kansen worden uitgewerkt in use-cases en geprioriteerd. Pak de meest relevante use-case en neem dit als uitgangspunt. Voordat je daadwerkelijk kan experimenteren met data is een cloud infrastructuur nodig, het fundament voor jouw platform. Natuurlijk kan je data ook direct in Excel inladen en een ‘Proof of concept’ uitwerken. Dit is een goede manier om zelf proefballonnen op te laten en zeker een goed startpunt voor analyses, maar zodra een initiatief rijp is om te delen met collega’s wil je de oplossing ook structureel goed borgen. Dan wil je een dataplatform tot je beschikking hebben. We sommen hier nog even kort de voordelen van een dataplatform op:

- Geautomatiseerd data verwerken en hiermee repeterend werk wegnemen
- Borgen ‘1 versie van de waarheid’
- Data uit meerdere bronsystemen samenvoegen en combineren
- Historie opbouwen over jouw data, wat weer vele nieuwe mogelijkheden van data-analyse met zich meebrengt (denk hier weer terug aan hoofdstuk 2: data gedreven adviseren).
- Grote rekenkracht om veel data te verwerken. Excel begint al snel te piepen en te kraken bij toename van data.
- Betere beveiliging van data

Terug naar dat het neerzetten van de basis infrastructuur, ook wel het fundament van jouw platform. Zonder technische achtergrond is dit behoorlijk lastig. Dus schroom niet om hier specialisten in te huren wanneer je deze kennis (nog) niet zelf in huis hebt. Net zoals bij het bouwen van een huis moet er een gedegen fundament zijn. Dit bepaalt in grote mate het succes van jouw dataplatform. Het zou dus ook een slecht advies zijn om hierop te beknibbelen.

Vanuit yellow arrow zien wij een cloud omgeving als een grote blokkendoos. De bouwblokken zijn de verschillende functionaliteiten in een cloud omgeving die je kan aanmaken en configureren. Bijvoorbeeld een opslag account of een database, maar ook infrastructuur onderdelen als een virtueel netwerk, sleutelkluisen om wachtwoorden in op te slaan en bijvoorbeeld een firewall. Het geheel van deze bouwblokken vormt jouw dataplatform. Het ontwerpen en inrichten van zo'n dataplatform wordt normaliter door een Data Architect gedaan. Natuurlijk zijn er ook andere rollen



die dit eventueel zouden kunnen, maar dan moeten deze specialisten wel een gedegen kennis hebben van cloud omgevingen en data architecturen. Hoe bepaal je nu welke bouwblokken in jouw situatie het beste passen?

Selecteren van de bouwblokken

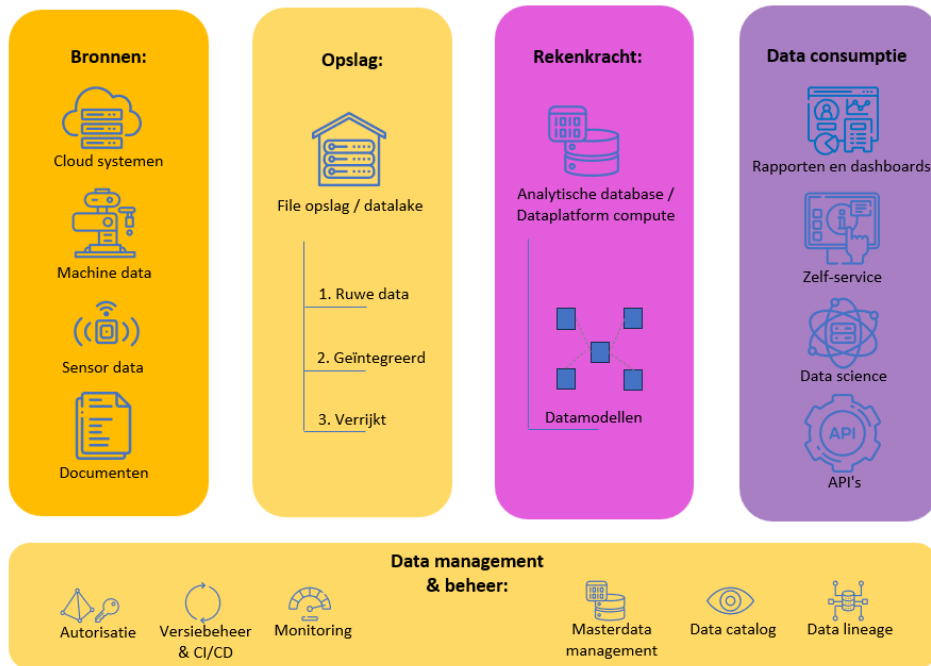
Hiervoor hebben we een aantal eenvoudige richtlijnen; Dingen makkelijker maken, daar houden we van bij yellow arrow.

- **Start klein;** Zoals we al een paar keer benoemd hebben: start klein, maar denk vooral groot. Ook bij het ontwerp van jouw dataplatform. Kies bouwblokken waar je met lage kosten kan instappen, maar die wel enorm schaalbaar zijn. Kies in de beginfase zo veel mogelijk voor 'pay per use' cloud consumptie. Je betaalt dan naar gebruik en als je bouwblokken niet gebruikt betaal je er ook niet voor. Houdt wel rekening mee dat je wel altijd betaald voor opslag van data, maar opslag is goedkoop. De rekenkracht om data te analyseren zal de meeste kosten met zich meebrengen. Door efficiënt gebruik van deze rekenkracht kan je kosten wel laag houden. Kijk vooral naar bouwblokken met rekenkracht welke automatisch aan gaan bij gebruik en ook weer uitgaan bij inactiviteit. Ook bouwblokken die automatisch op- en afschalen binnen een bepaalde bandbreedte zijn aan te raden. De 'pay per use' manier van consumeren biedt je flexibiliteit en kan je kosten besparen. Let wel op dat je de kosten goed blijft monitoren. Bouwblokken die niet automatisch uitgaan en vergeten worden om uit te zetten kunnen de kosten enorm laten oplopen.
- **Schaalbaarheid;** Dit betekent dat jouw dataplatform zich kan aanpassen aan veranderende omstandigheden zoals een toename (of afname) van de data of het gebruik ervan. Daarom is schaalbaarheid enorm belangrijk. Je wilt dat data binnen een acceptabele tijd beschikbaar is, maar je wilt ook grip houden op de kosten. Hier zit een tegenstrijdigheid in en zul je een juiste balans moeten zoeken. De meeste bouwblokken in de cloud zijn goed schaalbaar, daar zit immers ook een grote kracht van cloud omgevingen. Moderne analytische databases als Snowflake of aan dataplatform als Databricks of Microsoft Fabric zijn enorm schaalbaar, maar bieden dus ook de mogelijkheid om automatisch terug te schalen. Ook opslagaccounts als datalakes kunnen ontzettend hard groeien zonder dat het direct ten koste gaat van de performance van data-oplossingen.
- **Ontkoppelbaar;** Een belangrijk advies wanneer je een dataplatform in de cloud inricht is goed na te denken over ontkoppelpunten. Houdt data en rekenkracht zo veel mogelijk

gescheiden en zorg ook dat je verrijkte datamodellen weer zou kunnen opbouwen op basis van de ruwe data. Door hier bewust in jouw ontwerp rekening mee te houden is het ook makkelijker om bepaalde bouwblokken in te ruilen als ze niet de toegevoegde waarde brengen die je had verwacht of wanneer er simpelweg nieuwe bouwblokken komen die beter zijn. Waar het kan zorg dat technologie relatief eenvoudig in te wisselen is. Hierbij moeten we natuurlijk wel de kanttekening maken dat het altijd een bepaalde inspanning kost om bouwblokken te vervangen hoe goed alles ook ontworpen is.

- **Datakwaliteit;** Houdt rekening met de kwaliteit van data. Zorg altijd dat er plek is om de onbewerkte data op te slaan en dat je deze ook goed bewaard. Maak een tweede laag binnen de architectuur voor geïntegreerde en opgeschoonde data en een derde laag voor verrijkte data. Die integratie en verrijkings-laag kan je virtueel maken zodat je niet continue data aan het rond kopiëren bent. Zorg wel dat je altijd vastlegt hoe de data aangeleverd is, dus zonder aanpassingen. Dit noemen we ook wel de ruwe data. Hierdoor ben je in staat om te bepalen of data issues al in een bronsysteem ontstaan of door onjuiste logica op jouw dataplatform.
- **Security by design;** De beveiliging van jouw platform moet je ook niet lichtzinnig opvatten. Een hack of een data lek kan voor veel problemen zorgen. Denk aan reputatie schade of zelfs boetes vanuit toezichthouders. Het is niet meer acceptabel om hier vooraf niet over na te denken bij het maken van jouw ontwerp voor een dataplatform. Het grote voordeel is wel dat moderne cloud providers veel aandacht besteden aan beveiliging en er ook veel opties zijn om jouw data en jouw platform te beveiligen. Denk hierbij aan rol gebaseerde toegang, firewalls en detectie van onverwachte en schadelijke activiteiten.
- **Privacy by design:** Het zal je waarschijnlijk niet ontgaan zijn dat er ook veel gebeurd op het gebied van privacy. De wetgeving is de laatste jaren steeds strenger geworden met betrekking. Daarom is het, net zo als bij security, een vereiste om hier in jouw ontwerp over na te denken. Het is verstandig om risico's en eventuele schade uit te zetten tegenover de kosten van de te nemen maatregelen. Mocht je interessante data initiatieven bedacht hebben waarbij privacy gevoelige data nodig is dan moet je toetsen of er voor deze use-case(s) een grondslag is om privacy gevoelige informatie in een data platform op te slaan. Dit kan via een **Data Privacy Impact Assessment (DPIA)**. Dit kan als gevolg hebben dat je andere bouwblokken moet kiezen om de maatregelen toe te passen die vereist zijn.
- **Kennisniveau;** Het kennisniveau van collega's die het platform moeten onderhouden, maar ook de data bedrevenheid van de consumenten die platform gaan gebruiken moet je ook absoluut meenemen in keuze voor de bouwblokken. Wanneer je bijvoorbeeld kiest voor zelf-service oplossingen dan heeft een SQL gebaseerde technologie vaak een lagere instap. Oplossingen gebaseerd op programmeertalen als Python of Scala is meer specialistische kennis voor nodig. Ook de keuze voor SaaS, PaaS of IaaS kan afhangen van beschikbare kennis. Wil je volledig ontzorgd worden of wel je zelf meer kunnen configureren. Houdt er wel rekening mee dat wanneer je meer beheer naar jezelf toe trekt je hier ook de kennis voor in huis moet hebben.

- **Visualiseer de architectuur;** Om een dataplatform te ontwerpen is het goed om jouw ontwerp te visualiseren. Wij gebruiken daar normaliter onze dataplatform referentie architectuur voor. Dit is een uitwerking van alle functionaliteiten welke een dataplatform zou kunnen bevatten. We hebben een vereenvoudigde versie gemaakt van hoe zo'n visualisatie er uit kan zien. Deze visualisatie is nog zonder keuze voor specifieke bouwblokken, maar die kan je samen met jouw datateam en eventuele datapartner samen invullen. Bijvoorbeeld als opslag kan je voor een Azure datalake storage



(ADLS gen 2) kiezen en voor rekenkracht zou je Databricks of Microsoft Fabric kunnen inzetten. Met de richtlijnen die we hier hierboven beschreven hebben als leidraad kan je de bouwblokken naast elkaar zetten om keuzes te maken.

Deze richtlijnen geven jou houvast bij het maken van keuzes. We hebben hierboven de meest relevante richtlijnen benoemd. Natuurlijk zijn er nog meer van deze richtlijnen te bedenken. In IT-jargon zul je ook vaak de term 'guiding principles' horen vallen voor dit soort fundamentele richtlijnen. Deze guiding principles zijn geen wetten, maar je moet als ontwerper wel zwaarwegende redenen aandragen om hiervan af te wijken. Het is verstandig om deze guiding principles uit te werken en de implicaties te beschrijven. Maak er eventueel een one-pager van die je op de muur (of ergens op een digitale werkplek, als er veel remote gewerkt wordt) hangt bij de domein teams en het centrale data team, zodat iedereen weet wat deze richtlijnen zijn en iedereen dezelfde taal gaat spreken. TOGAF is een architectuur framework wat hierbij kan helpen. Op de website van TOGAF kan je bijvoorbeeld veel documentatie terugvinden hoe je een data architectuur en bijbehorende guiding principles uitwerkt. Zie TOGAF als een grote gereedschapskist waar jij de juiste tools uit kan gebruiken.

Wanneer je een blauwdruk voor jouw dataplatform hebt gemaakt kan je de bouwblokken daadwerkelijk uitrollen in de cloud omgeving naar keuze en kan je aan de slag met de data use-case(s). Ga vooral beginnen, experimenteer en leer. Laat je goed adviseren als je zelf niet de kennis in huis hebt en leer stapje voor stapje als organisatie en als individu om data bedreven te worden. En onthoudt; Een kasteel wordt niet in 1 dag gebouwd, maar met de kennis van dit whitepaper en een goede mindset zijn wij ervan overtuigd dat het voor elk accountskantoor mogelijk is om je eigen kasteel te bouwen. De data gedreven accountant is nu aan zet!



Conclusie

De wereld van de accountant veranderd en dat biedt nieuwe kansen. Wel is het zo dat je verandering niet uit de weg moet gaan want anders loop je het risico jezelf minder relevant te maken. Pak de data regie stevig in handen. Ga zelf aan de slag met data en verzamel data in een eigen analyse omgeving. Zoals we in deze whitepaper toegelicht hebben biedt dit veel voordelen zoals flexibiliteit en efficiëntie. Daarnaast hou je zelf de regie over jouw data en word je niet 'gegijzeld' door softwareleveranciers.

Wil je aan de slag met data maar heb je hulp nodig bij het verkennen van kansen, het uitwerken van een use-case of het ontwikkelen van een dataproduct dan helpen wij daar graag bij. Neem gerust contact met ons op: [Contact | yellow arrow](#)

yellow arrow

**Marconiweg 22a
8071RA Nunspeet**

hallo@yellowarrow.nl

085 273 8313

yellowarrow.nl

The background consists of several overlapping geometric shapes. A large yellow arrow points from the top-left towards the center-right. A dark blue arrow points from the top-right towards the center-left. A light pink arrow points from the bottom-left towards the center-right. The overall composition is dynamic and modern.