



Data Management Croonwolter&dros

Afstudeerverslag

AUTEUR	Dennis Loos
STUDENTNUMBER	S1096946
SCHOOL	Hogeschool Leiden
FACULTEIT	Informatica – Business Data & Management
STARTDATUM	06-02-2019
EINDDATUM	31-05-2019
ORGANISATIE	CroonWolter&Dros
AFDELING	STAF-ICT-Informatie Management
EXAMINATOR HOGESCHOOL LEIDEN	James Walls
BEGELEIDER HOGESCHOOL LEIDEN	Mischa Barthel
BEGELEIDER CROONWOLTER&DROS	Rob Kabel
VERSION NUMBER	3.0
VERSION DATE	07-06-2019
VERSION TYPE	Final



Voorwoord

Voor u ligt de scriptie 'Data Management Croonwolter&dros'. Het onderzoek is uitgevoerd in opdracht van de inkoopafdeling binnen de organisatie Croonwolter&dros. Deze scriptie is geschreven in het kader van mijn afstuderen aan de opleiding Informatica – Business Data & Management aan de Hogeschool Leiden.

Samen met mijn mede-afstudeerder Sike van Dongen van de Haagse Hogeschool hebben we het onderzoek uitgevoerd. Ik heb het onderdeel datakwaliteit management uitgevoerd en Sike heeft nieuwe rapportages ontwikkeld voor de inkoopafdeling.

Bij deze wil ik graag mijn begeleider Rob Kabel bedanken voor de ondersteuning en uitzonderlijk goede begeleiding tijdens dit traject. Rob heeft altijd tijd gemaakt om feedback of ondersteuning te geven en om mee te denken met het onderzoek. Tijdens dit onderzoek heb ik inhoudelijk ondersteuning gekregen van Albert Huttenga en Eltjo Renken. Zonder hun ondersteuning zou ik dit onderzoek niet succesvol kunnen afronden.

Ook wil ik de collega's van de BI afdeling bedanken voor de ondersteuning en fijne samenwerking. Ik heb veelvuldig uitleg en feedback van hen gekregen over de organisatie en het onderzoek. Zij hebben mij goed opgenomen in het team, waardoor het fijn was om op kantoor te werken.

Dennis Loos

Leiden, 11 juni 2019

Managementsamenvatting

Voor middel tot grote organisaties is data management een essentieel bedrijfsproces. Beslissingen worden gemaakt op basis van data, deze wordt ingevoerd vanuit verschillende locaties en afdelingen. Het is van belang dat afspraken gemaakt worden over hoe er met data wordt omgegaan binnen een organisatie. Dit hoort op een gestructureerde manier uitgevoerd te worden, waarbij definities en begrippen hetzelfde betekenen binnen de hele organisatie.

Dit onderzoek heeft op basis van een functioneel kader en kwalitatief & kwantitatief onderzoek een proof of concept gemaakt om data management te integreren binnen Croonwolder&dros. Hiervoor is de volgende onderzoeksvraag opgesteld: *Welke maatregelen dienen gemaakt te worden om de datakwaliteit te waarborgen wanneer de data toegankelijkheid verhoogd wordt.* Datakwaliteit staat hierbij centraal.

Om een antwoord te kunnen geven op de hoofdvraag, is deze opgedeeld in 9 deelvragen. In deze deelvragen worden de wensen en eisen achterhaald van de opdrachtgever en de stakeholders. Op basis van een raamwerk is een datakwaliteit management tool ontwikkeld die een gestandaardiseerde score geeft. Deze score vertelt in hoeverre de data volgens de normen van de organisatie correct is.

Er zijn verschillende mogelijkheden onderzocht om de oplossing in te ontwikkelen en er is een proof of concept gerealiseerd die in productie wordt genomen door de organisatie. Vervolgens wordt er een conclusie getrokken met aanbevelingen op basis van de resultaten.

Versiebeheer

Versie	Datum	Status	Opmerkingen
0.1	20-04-2019	Ontwikkel	Opzet document, Organisatie beschrijving, stakeholders, methodologie
0.2	23-04-2019	Ontwikkel	A-Competenties, B-Competentie
1.0	23-04-2019	Concept	Omzetten naar concept
1.1	16-05-2019	Concept	Feedback verwerkt, kwantitatief onderzoek & doelstelling toegevoegd.
1.2	28-05-2019	Concept	Conclusie en Reflectie
1.3	29-05-2019	Concept	Inhoudelijke wijzigingen
2.0	01-06-2019	Pre-Final	-
2.1	04-06-2019	Pre-Final	Feedback stagebegeleiders verwerkt
3.0	07-06-2019	Final	Competenties gemodificeerd



Inhoudsopgave

Voorwoord	2
Managementsamenvatting	3
Versiebeheer	4
Tabellen	7
Figuren.....	8
Begrippenlijst.....	9
1 Inleiding	10
1.1 Aanleiding.....	10
1.2 Probleemstelling.....	11
1.3 Doelstelling.....	12
2 Beschrijving organisatie.....	13
2.1 De organisatie.....	13
2.2 Organisatiestructuur	14
2.3 Gegevens	15
2.4 Begeleiders	16
2.5 Mede-afstudeerder	17
3 Methodologie	18
3.1 Kwalitatief en kwantitatief onderzoek	18
3.2 SCRUM-methode.....	19
3.3 MoSCoW	21
3.4 DAMA DMBOK2 Raamwerk.....	21
3.5 Microsoft DevOps.....	22
4 Onderzoeksvragen.....	24
5 Onderzoeksresultaten	25
5.1.1 Welke KPI's en requirements zijn gedefinieerd binnen de inkoopafdeling?	25
5.1.2 Wie zijn de stakeholders en welk belang hebben zij bij het project?	27
5.1.3 Wat doet de organisatie aan data management?.....	29
5.1.4 Welke invloed heeft de data migratie naar Microsoft Azure op het project?	30
5.1.5 Welke aspecten van datamanagement hebben een impact op dit project?	34
5.1.6 Hoe wordt de datakwaliteit getoetst?	37
5.1.7 Welke tools kunnen gebruikt worden voor datakwaliteit management?	48
5.1.8 Welke opmerkelijke datakwaliteit gerelateerde bevindingen zijn er gevonden?.....	53
5.1.9 Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?.....	55
6 Conclusie	58



6.1	Discussie	59
6.2	Aanbevelingen	60
7	Competenties	61
7.1	A-competenties	61
7.1.1	Onderzoek	61
7.1.2	Leren leren	61
7.1.3	Professioneel werken	62
7.1.4	Innovatie	62
7.2	B-Competenties	62
7.2.1	Bedrijfsprocessen analyseren	62
7.2.2	Infrastructuur analyseren	63
7.2.3	Bedrijfsprocessen ontwerpen	63
8	Reflectie	65
8.1	Planning wijzigingen	65
8.2	Evaluatie	66
9	Bibliografie	68
10	Bijlagen	70
10.1	Bijlage 1 Requirements en KPI's in DevOps	70
10.2	Bijlage 2 Oude dashboard inkoopafdeling	72
10.3	Bijlage 3 Workshop stakeholders presentatie	73
10.4	Bijlage 4 Datakwaliteit SQL scripts	80
10.5	Bijlage 5 meetresultaten Excel en Power BI	83
10.6	Bijlage 6 meetresultaten SQL Server	84
10.7	Bijlage 7 DAX Statements formule gestandaardiseerd datakwaliteit cijfer	85
10.8	Bijlage 8 Blauwdruk schets Datakwaliteit dashboard	86
10.9	Bijlage 9 planning	87
10.10	Bijlage 10 management presentatie	90
10.10.1	Demo Croonwouter&dros inkoopdashboard	93
10.11	Bijlage 11 meeting factuurregels zonder kostenplaats	94
10.12	Bijlage 12 Assessment data mart crediteuren	96
10.13	Bijlage 13 Datakwaliteit Meetinstrument	98
10.14	Bijlage 14 Plan van Aanpak	127



Tabellen

Tabel 1 Informatie Croonwolver&dros vestigingen.....	15
Tabel 2 Croonwolver&dros begeleider en opdrachtgever	16
Tabel 3 Hogeschool Leiden begeleiders	17
Tabel 4 Mede-afstudeerder informatie.....	17
Tabel 5 DevOps begrippen	22
Tabel 6 Key performance indicatoren	26
Tabel 7 stakeholderanalyse.....	27
Tabel 8 data gewichten	40
Tabel 9 voorbeeld formule	40
Tabel 10 datakwaliteit aspecten meta data	41
Tabel 11 _dataerror.....	44
Tabel 12 datakwaliteit tool requirements.....	48
Tabel 13 vergelijkingsonderzoek Dataladder	49
Tabel 14 Technische Unie verschil	54
Tabel 15 uitleg dashboard.....	56
Tabel 16 data assessment opmerkingen.....	97
Tabel 17 data assessment aanpassingentabel	97

Figuren

Figuur 1 Organigram organisatiestructuur	14
Figuur 2 Organigram Business Information Team.....	15
Figuur 3 scrum sprint framework.....	20
Figuur 4 work items DevOps	23
Figuur 5 ETL proces on-premise	31
Figuur 6 ETL-proces Azure	32
Figuur 7 Data Management drie domeinen	34
Figuur 8 Datakwaliteit formule	40
Figuur 11 vergelijking Excel en SQL Server meting.....	42
Figuur 12 relaties datakwaliteit tabellen.....	45
Figuur 13 uitleg inhoud tabellen leverancier data	45
Figuur 14 validatie formule datakwaliteit	47
Figuur 10 Power BI Data Profiling	50
Figuur 15 omzet per project.....	53
Figuur 16 bestelorder scherm acto	53
Figuur 17 projectnummers en kostenplaats registratie ACTO	54
Figuur 18 Data Matching 1	54
Figuur 19 Data Matching 2	54
Figuur 20 DQ leverancier dashboard.....	56
Figuur 21 DQ definitie pagina.....	57
Figuur 22 requirements Excel.....	70
Figuur 23 KPI's Excel	70
Figuur 24 DevOps product backlog	70
Figuur 25 Dev ops user stories	71
Figuur 26 Startpagina	72
Figuur 27 Leverancierskaart	72
Figuur 28 DQ dashboard schets	86

Begrippenlijst

Begrip	Betekenis
Key Performance Indicator (KPI)	Variabelen om de prestaties van een bedrijfsproces te analyseren. KPI's worden gebruikt om het succes van een organisatie te kunnen meten.
Acto	Leverancier van het datawarehouse en het ERP-systeem.
Proof of concept	Een methode om door middel van een werkend product te demonstreren of een bepaalde IT-oplossing geschikt is voor een bepaald doel.
Bonustarget	Wanneer een bepaalde hoeveelheid in euro's wordt besteld bij een leverancier, kan Croonwolver&dros een X percentage korting krijgen van de leverancier.
Citrix	Een softwareprogramma die vereist is om toegang te krijgen tot Acto, dit gaat via een virtual machine.
Role Based Access Control	Is een methode waarmee op een effectieve en efficiënte wijze toegangscontrole voor informatiesystemen kan worden ingericht.
On-premise	Software en servers dat lokaal geïnstalleerd is en draait op de computers van de persoon of organisatie.
API	Application Programming Interface is een programma dat kan communiceren met andere programma's of onderdelen.
View	Een virtuele tabel. De data die in de view staat wordt bepaald aan de hand van de query die de data uit de tabel haalt.
Staging area	Tijdelijk opslaggebied tussen de databronnen en een datawarehouse.
Master database	Kritische data voor de organisatie of bedrijfsproces.
ETL	Extract Transform Load: Het proces dat de data ophaalt, transformeert en in het datawarehouse laadt.
Netwerkschijf	Een opslagmedium dat op het netwerk aangesloten is en gebruikmaakt van het TCP/IP-protocol voor dataoverdracht.
Bedrijfsproces	Activiteiten om een product of dienst te leveren.
SQL Query	Een opdracht in SQL taal die communiceert met een database om een bepaalde actie uit te voeren.
Surrogate key	Een surrogaat sleutel in een database is een unieke identificatie voor een entiteit in de gemodelleerde wereld of een object in de database.
DAX Statement	DAX staat voor Data Analysis Expressions en is een formuletaal in Power BI.
Data Factory	een service die is gebouwd voor alle gegevensintegratiebehoeften en vaardigheidsniveaus. Data Factory ondersteunt ETL- en ELT-processen en biedt de mogelijkheid om zonder code in de visuele omgeving te werken.
Basware	Basware Corporation is een Fins softwarebedrijf dat bedrijfssoftware verkoopt voor financiële processen, purchase to pay en financieel beheer.

1 Inleiding

1.1 Aanleiding

Ieder jaar vergroot de hoeveelheid gegenereerde data exponentieel. Dit komt mede door Internet of Things (IoT) en digitalisering. In 2020 wordt het 5G netwerk uitgerold, waardoor het genereren van data alleen nog maar meer wordt. Auto's kunnen met elkaar communiceren, Alle onderdelen in gebouwen kunnen verbonden worden met elkaar via een netwerk van sensoren, internet en andere technische apparaten.

Sinds 1 januari 2017 zijn Croon en Wolter & Dros gefuseerd tot Croonwolter&dros. Croon was gespecialiseerd in elektrotechniek en automatisering. Wolter & Dros was gespecialiseerd in werktuigbouwkunde en meet- en regeltechniek. De fusie betekende ook dat de bedrijfsprocessen, werkwijzen en de afspraken volgens een vaste norm moeten worden vastgelegd.

Hoge datakwaliteit is afhankelijk van structuur, precisie, procesmatig werken, de business en definities. Croonwolter&dros heeft verschillende administraties die data vastleggen in het datawarehouse. Met meerdere databronnen die de data invoeren groeit het risico op lage datakwaliteit. Maar wat is datakwaliteit nu precies? "datakwaliteit is de mate waarin data geschikt is voor het doel waarvoor ze gebruikt wordt".

Het risico van het ontbreken van datamanagement kan op ten duur voor grote problemen zorgen. Croonwolter&dros stuurt op basis van data. Als de data die weergegeven wordt niet overeenkomt met de werkelijkheid, kan dit resulteren in het nemen van verkeerde beslissingen. Onderhandelingen met klanten zijn gebaseerd op verkeerde informatie, of bonusafspraken worden niet behaald door het ontbreken van data. Het belang van datamanagement zal met de tijd alleen maar groeien en organisaties moeten hier met mee groeien, anders kunnen de kosten oplopen tot in de miljoenen.

Om te voorkomen dat de datakwaliteit verslechterd is er datakwaliteit management nodig. Met datakwaliteit management is een goede samenwerking tussen de business en IT vereist. Tools dragen bij aan het verhogen van de datakwaliteit, maar de business moet ook het belang ervan inzien en hun bijdrage.

1.2 Probleemstelling

Croonwolter&dros verwerkt meer dan 200.000 facturen per jaar en de inkoopafdeling behandelt 80% van de uitgaven van de organisatie. Inkoopdata is één van de basispilaren voor inzichtelijkheid en control van business performance en bewaking. Voor een goed bedrijfsbeleid en financieel beheer is het van belang dat de inkoop gecontroleerd en inzichtelijk is. Betrouwbare en tijdige data is van fundamenteel belang voor het beheren van processen en het behouden van overzicht over de verplichtingen. Dit is van toepassing op zowel operationeel als op strategisch niveau.

Naast het beschikbaar zijn van de data is het ook essentieel om te allen tijde te kunnen vaststellen hoe betrouwbaar de data is. De afdelingen hebben voor dezelfde situaties andere begrippen. Een leverancier bij de inkoopafdeling, is een crediteur binnen de financiën afdeling. Dit beïnvloedt de data integriteit. Het systeem binnen Croonwolter&dros heeft hier geen vast definities voor gedefinieerd.

Centrale dashboards die de inkoopafdeling ondersteunen bij het nemen van de juiste beslissingen en ook inzicht geven in de datakwaliteit zijn van invloed op de kostenbeheersing. Het huidige dashboard van de inkoopafdeling bevat geen real-time overzichten van de KPI's.

De inkoopafdeling werkt met bestelorders en de crediteuren afdeling werkt met factuurorders. De bestelorders bevatten de aantal producten/artikelen die besteld zijn bij één of meerdere leveranciers. De factuurorders bevatten de kosten van de bestelorders die geplaatst zijn door de inkoopafdeling. De twee datasets hebben een relatie met elkaar, elk factuur is berekend op de bestelorders. In het huidige datawarehouse worden de datasets niet gecombineerd met elkaar. Hierdoor kan Croonwolter&dros de omzet niet betrouwbaar berekenen en de huidige status van processen inzien.

1.3 Doelstelling

Het doel van dit project is om het huidige inkoopproces te optimaliseren. Datakwaliteit is een opkomend probleem bij de organisatie en speelt een essentiële rol bij het nemen van beslissingen op strategische en operationele niveaus.

De output van dit project is een proof of concept waarin een klein onderdeel van het datawarehouse op datakwaliteit gemeten is. Dit proof of concept bestaat uit een theoretisch kader, een advies rapport, en een product. Het product bestaat uit resultaten van de meting en een datakwaliteit dashboard. Het datakwaliteit dashboard weergeeft de huidige datakwaliteit binnen de organisatie.

In de huidige situatie is er geen structurele data kwaliteitsmanagement binnen de organisatie. Zonder een indicatie te hebben in welke hoedanigheid de data van hoge kwaliteit is, kunnen geen betrouwbare beslissingen genomen worden op basis van de data. Data kwaliteit management is van belang om de kwaliteit te waarborgen en om met zekerheid te kunnen stellen dat de data overeenkomt met de werkelijkheid.

Met het ontwikkelen van een datakwaliteit meetinstrument kunnen de dashboard resultaten voor de inkoopafdeling worden gevalideerd met het datakwaliteit meetinstrument. Het meetinstrument geeft een gestandaardiseerde waarde terug over de hoedanigheid van de data. Het product draagt bij aan:

- Minder tijd verloren aan het valideren of de gegevens kloppen met de werkelijkheid.
- Betere onderhandelingsresultaten met leveranciers door meer inzicht te hebben in de huidige leveranciersstatus.
- Inzicht krijgen op de huidige uitgaven en de bonustargets.
- Meetbaar maken van de kwaliteit van de data.
- Foutieve data assets kunnen identificeren en modificeren.
- Data beheerders kunnen zelf de foutieve data identificeren en modificeren.

De data beheerders/data stewards zijn verantwoordelijk voor het managen van de data. Met het product kan iedere afdeling hun eigen datakwaliteit beheren en verbeteren.

2 Beschrijving organisatie

2.1 De organisatie

Croonwolter&dros loopt al meer dan 140 jaar voorop bij nieuwe ontwikkelingen. Ook in deze tijd, waarin de rol van techniek steeds belangrijker wordt, is Croonwolter&dros een aanjager van nieuwe technologische ontwikkelingen. Croonwolter&dros heeft de kennis, de ervaring en het vakmanschap om intelligente systemen en integrale, duurzame oplossingen te bedenken, te ontwikkelen, te implementeren, te beheren en te onderhouden.

Vanuit het motto 'Intelligentie door Technologie' draagt Croonwolter&dros met intelligente technologische elektrotechnische-, werktuigbouwkundige- en automatiserings- en informatisering oplossingen bij aan het duurzaam presteren van haar klanten. Met deze intelligente, zelfdenkende en -lerende systemen worden processen in of van objecten zoals schepen, tunnels, gebouwen en fabrieken efficiënter en effectiever.

Croonwolter&dros komt voort uit de fusie van Croon en Wolter & Dros. Croon was gespecialiseerd in elektrotechniek en automatisering. Wolter & Dros was gespecialiseerd in werktuigbouwkunde en meet- en regeltechniek. Met een rijke historie van 140 jaar Croon en 141 jaar Wolter & Dros zijn de ondernemingen samen goed voor meer dan 280 jaar ervaring in de techniek en opereren nu onder de merknaam Croonwolter&dros.

Met een omzet van 550 miljoen en 3.000 medewerkers is Croonwolter&dros het grootste in elektrotechniek, werktuigbouwkunde, automatisering en informatisering gespecialiseerde bedrijf van Nederland. Croonwolter&dros is actief in de volgende markten: Utiliteitsbouw, Industrie, Infra en Marine & Offshore en heeft vestigingen in heel Nederland. Croonwolter&dros is onderdeel van de TBI-holding.

De afdeling Croonwolter&dros STAF-Inkoop-Support is een subafdeling van de afdeling Inkoop. De gehele inkoopafdeling bestaat uit ca. 38 medewerkers en de sub afdeling bestaat uit 10 medewerkers. De afdeling Croonwolter&dros STAF-Inkoop-Support geeft ondersteuning aan het inkoopproces en afdeling inkoop in breedste zin. De primaire processen binnen deze afdeling is o.a. zorgdragen voor:

- Analyses van het inkoopproces
- Het beschikbaar stellen van management rapportages.
- Datamanagement m.b.t. inkoopproces
- Procesoptimalisatie

De informatievoorzieningen die gebruikt worden binnen de afdeling Croonwolter&dros STAF-Inkoop-Support zijn datawarehousing, Power BI, Qlikview en een ERP (Enterprise Resource Planning) -pakket, Acto. De bedrijfscultuur in relatie tot de informatievoorziening is traditioneel/ conservatief.

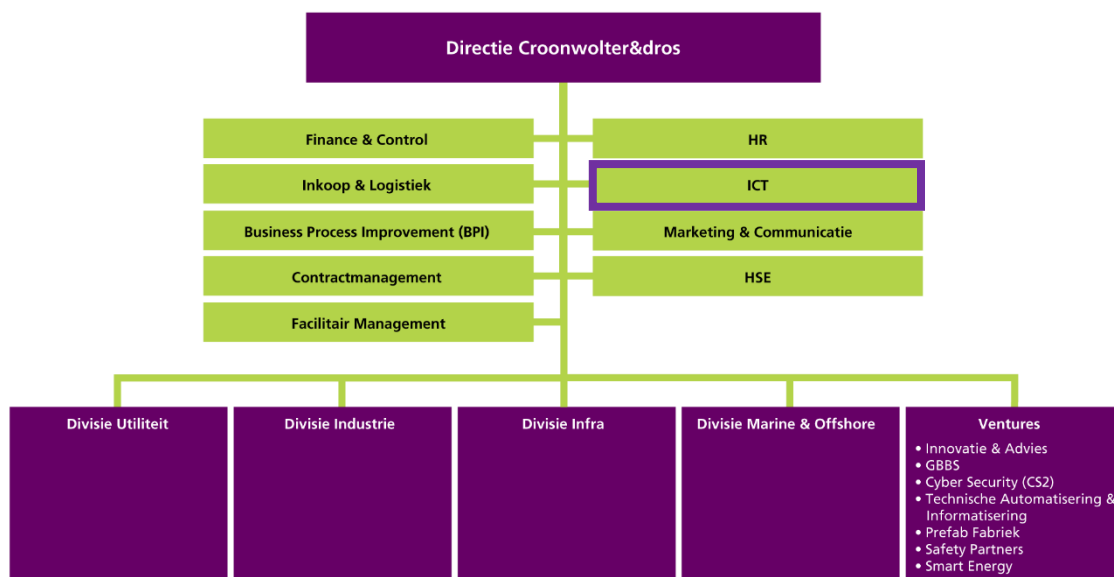
2.2 Organisatiestructuur

Onderstaand figuur toont de lijn-staf-organisatie structuur van Croonwolder&dros. Alle staffuncties ondersteunen de vijf onderliggende divisies. De divisie “Ventures” zijn kleine ondernemingen die ieder specifieke bijdragen leveren aan het functioneren van Croonwolder&dros.

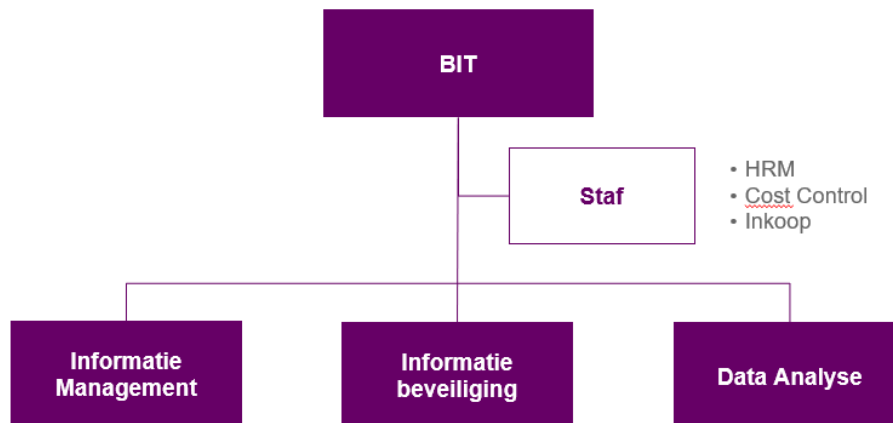
Het project is in opdracht van de inkoopafdeling. Het Business Intelligence Team (BI) voert de opdracht uit. BI bevindt zich onder de ICT-afdeling. De afstudeerders zitten op de BI afdeling. In Figuur 2 is de organigram van BI weergegeven. De afstudeerders werken als data analisten voor de organisatie. Vanuit de BI afdeling krijgen de afstudeerders ondersteuning bij technische vragen of over vragen over de business.

Met het komen van een Shared Service Center (SSC) is Servicedesk & Beheer verhuist naar een andere locatie. BI adviseert de business bij vraagstukken tussen business en IT en ondersteund hierbij bij het implementeren van oplossingen. BI bewaakt het applicatielandschap, informatiestromen en security-aspecten. Vanuit BI wordt de organisatie voorzien van rapporten en dashboards.

Het SSC is verantwoordelijk voor de operationele dienstverlening m.b.t. ICT. Deze operationele dienstverlening omvat o.a. beheer van de technische infrastructuur, beheer van werkplekken, printers en de ondersteuning van generieke applicaties.



Figuur 1 Organigram organisatiestructuur



Figuur 2 Organigram Business Information Team

2.3 Gegevens

Croonwouter&dros heeft verschillende vestigingen in het land. De opdracht wordt uitgevoerd door twee stagiaires. Om te verzekeren dat beiden afstudeerders de competenties kunnen aantonen, is de opdracht verdeeld in twee verschillende onderdelen. Dit wordt verder uitgelegd in het hoofdstuk [Mede-afstudeerder](#).

De afstudeerders werken op twee verschillende locaties, in Rotterdam en Amersfoort. Hier is voor gekozen omdat de medewerkers van de afdeling Staf-Inkoop-Support verspreid zitten over deze twee locaties. Beiden afstudeerders werken onder de Business Intelligence afdeling van Croonwouter&dros. De BI-afdeling implementeert de opdracht voor de afdeling Staf-Inkoop-Support.

Vestiging	Rotterdam	Amersfoort
Adres	Marten Meesweg 25	Amsterdamseweg 53
Postcode	3068 AV	3812 RP
Plaats	Rotterdam	Amersfoort
Telefoon	088 923 33 44	088 923 61 61

Tabel 1 Informatie Croonwouter&dros vestigingen

2.4 Begeleiders

De begeleider voor de afstudeerder binnen de organisatie is Rob Kabel. Rob is het aanspreekpunt wanneer er vragen, problemen en/of opmerkingen zijn. De eerste weken plant Rob (introductie) meetings in om de afdeling en de manier van werken binnen de organisatie te leren kennen. Tijdens de afstudeeropdracht blijft Rob betrokken bij de opdracht, om te controleren of de scope van de opdracht niet afwijkt en ondersteund de afstudeerder indien nodig.

De opdrachtgever van de afstudeeropdracht is Mathieu Koomans v.d. Dries. Mathieu geeft informatie over de opdracht en ondersteuning indien nodig. Aan het einde van de afstudeerperiode beoordeelt Mathieu de resultaten. Mathieu wordt tussentijds op de hoogte gehouden over de voortgang van de opdracht.

Achternaam opdrachtgever:	Koomans v.d. Dries
Voorletters opdrachtgever:	M.L. (Mathieu)
Titulatuur opdrachtgever:	Dhr.
Functie opdrachtgever:	Manager Procurement Support
Doorkiesnummer opdrachtgever:	+31 88 923 34 19
Email opdrachtgever:	mathieu.koomansv.d.dries@croonwolderendros.nl
Achternaam bedrijfsmentor:	Kabel
Voorletters bedrijfsmentor:	R.M.A. (Rob)
Titulatuur bedrijfsmentor:	Dhr.
Functie bedrijfsmentor:	Manager Business Intelligence
Mobiel bedrijfsmentor:	+31 6 12 91 03 94
Email bedrijfsmentor:	Rob.kabel@croonwolderendros.nl

Tabel 2 Croonwolder&dros begeleider en opdrachtgever

De docentbegeleider van de Hogeschool Leiden is Mischa Barthel. Mischa is het aanspreekpunt van school als er problemen en/of opmerkingen zijn. De begeleider komt minimaal één keer langs bij de organisatie. Dit (eerste) bezoek is in de tweede of derde week van de afstudeerperiode. Aan de hand van het bezoek in de tweede of derde week geeft de docentbegeleider een GO/NO-GO voor het afstudeerplan.

De examinerator van de Hogeschool Leiden is James Walls. De examinerator stelt het eindcijfer vast vanuit de Hogeschool Leiden. Dit wordt gedaan samen met de docentbegeleider en een externe beoordelaar aan de hand van het op te leveren afstudeerdossier, de verdediging en het advies van de bedrijfsbegeleider.

Achternaam examinerator:	Walls
Voorletters examinerator:	J.
Titulatuur examinerator:	Dhr.
Functie examinerator:	Docent informatica
Telefoonnummer examinerator:	+31611585667
Email examinerator:	Walls.j@hsleiden.nl
Achternaam docentbegeleider:	Barthel
Voorletters docentbegeleider:	M.
Titulatuur docentbegeleider:	Dhr.
Functie docentbegeleider:	Docent informatica
Mobiel docentbegeleider:	+31648133867
Email docentbegeleider:	barthel.m@hsleiden.nl

Tabel 3 Hogeschool Leiden begeleiders

De afstudeerder is verantwoordelijk voor het afstudeerproject en heeft zelf de opdracht geformuleerd en verzorgt. De afstudeerder is eigenaar van het afstudeerdossier. Tijdens het afstudeertraject kan de afstudeerder gebruik maken van de ondersteuning van Croonwolter&dros en Motion10 (Motion10, 2019). De afstudeerder is via Motion10 bij Croonwolter&dros terecht gekomen. Tijdens het afstudeertraject is Motion10 een informele ondersteuning voor de afstudeerders. Hier kunnen de studenten advies of ondersteuning vragen als zij vastlopen in het proces.

2.5 Mede-afstudeerder

Aan deze opdracht wordt samengewerkt met een andere afstudeerder. De opdracht is onderverdeeld in twee onderdelen. Mijn focus ligt bij de datakwaliteit en de data toegankelijkheid. Mijn verantwoordelijkheid is om een datakwaliteit tool en een adviesrapport op te stellen waarmee Croonwolter&dros datakwaliteit management kan implementeren. De andere component van de opdracht is de data visualisatie. Mijn mede-afstudeerder richt zich op de User Experience (UX) kant van het project en het ontwikkelen van dashboards. Dit omvat het ontwerpen van de rapportages en dashboards inclusief de functionaliteiten in Power BI.

Naam	S. van Dongen
School	Haagse Hogeschool
Opleiding	Business IT & Management

Tabel 4 Mede-afstudeerder informatie

3 Methodologie

In dit hoofdstuk worden de gebruikte methoden beschreven die binnen het project zijn gebruikt. Vragen als wat voor onderzoek er is gedaan, hoe de data en requirements zijn verzameld, wat voor methoden er zijn toegepast en hoe er tot de resultaten is gekomen, worden in dit hoofdstuk beschreven.

3.1 Kwalitatief en kwantitatief onderzoek

Er is een kwalitatief en kwantitatief onderzoek uitgevoerd om een meetinstrument te ontwikkelen voor Croonwolter&dros.

Kwalitatief onderzoek

Kwalitatief onderzoek is toegepast om interpretaties, definities en gedachten van de betrokken medewerkers te verzamelen. De gebruikers van de inkoopdata hebben verschillende functies binnen de organisatie en ieder individu gebruikt de data voor zijn eigen doeleinde.

Om te achterhalen wat de verschillen zijn tussen de gebruikers, wat hun meningen zijn over de huidige status en wat zij belangrijk vinden dat verbeterd moet worden, is er een workshop gehouden met vijf gebruikers (Zie *Bijlage 3 Workshop stakeholders presentatie*). De vijf inkopers zijn representatief voor hun functiegebied. Tijdens de workshop is er met post-its gebruik gemaakt van de MoSCoW methode om de wensen en eisen van de stakeholders te verzamelen en prioriteren.

Om meer informatie te verzamelen omtrent de definities die gelden binnen de organisatie, zijn verschillende interviews afgenomen en meetings georganiseerd met de stakeholders en business experts. Tijdens deze sessies zijn vraagstukken beantwoord over de betekenis van bepaalde data entiteiten en business gerelateerde onderwerpen. Óók zijn er meetings geweest waarin de proces- en data flow uitgelegd en beschreven zijn om de werkwijze en de architectuur beter te begrijpen.

Tijdens meerdere sessies met de leverancier van de data marts is er een assessment gedaan op de data. Het document (*Bijlage 12 Assessment data mart crediteuren*) bevat een assessment op één data mart. Hierin staat beschreven wat de bevindingen zijn geweest en welke prioriteiten ze hebben. De resultaten zijn opgestuurd naar Acto met het verzoek om dit door te voeren in het systeem.

De resultaten van de workshop, meetings en de interviews zijn verwerkt in de Microsoft DevOps tool. Deze online tool wordt gebruikt om met de *SCRUM-methode* te werken.

Kwantitatief onderzoek

Een kwantitatief onderzoek is toegepast op het theoretisch kader en het implementeren van het datakwaliteit-meetinstrument. Het theoretisch kader omvat alle onderdelen die van belang zijn bij het opzetten van data management voor Croonwolter&dros. Uit het theoretisch is er dieper ingezoomd op de domeinen die van belang zijn voor het uitvoeren van deze opdracht.

Het onderdeel datakwaliteit management uit het theoretisch kader wordt uitgevoerd in deze opdracht. De data wordt verzameld uit het datawarehouse van Croonwolter&dros. Op basis van de gedefinieerde data rules, wordt de datakwaliteit meting uitgevoerd over de tabellen. Met een formule wordt het gewogen gemiddelde berekend en retourneert de formule één gestandaardiseerde eindscore.

3.2 SCRUM-methode

Het realiseren van de opdracht wordt uitgevoerd volgens de SCRUM-methode. Wat is SCRUM precies? SCRUM is een flexibele manier van werken, waarbij veel interactie is tussen de projectleden. Er gaat minder tijd in de voorbereiding zitten en problemen kunnen snel opgelost worden.

De reden dat er voor SCRUM is gekozen, is omdat het nog niet duidelijk is hoe we bepaalde activiteiten aanpakken. Gedurende het project kunnen er veranderingen optreden, dit kan uit externe- of interne afhankelijkheden komen. Met deze onzekerheden is SCRUM een gepaste methode omdat er geen planning van begin tot eind is gemaakt. Gedurende het project kunnen er veranderingen plaatsvinden, met SCRUM kun je hier snel op inspelen en de impact van problemen minimaliseren. Met sprints van twee weken wordt er gestructureerd naar het eindresultaat gewerkt. In de [Bijlage 14 Plan van Aanpak](#) staat de planning met de activiteiten die uitgevoerd zijn in iedere iteratie.

SCRUM-tool

De tool die wordt gebruikt om via de SCRUM methode te werken is DevOps. In het hoofdstuk [Microsoft DevOps](#) wordt de tool beschreven.

SCRUM-team

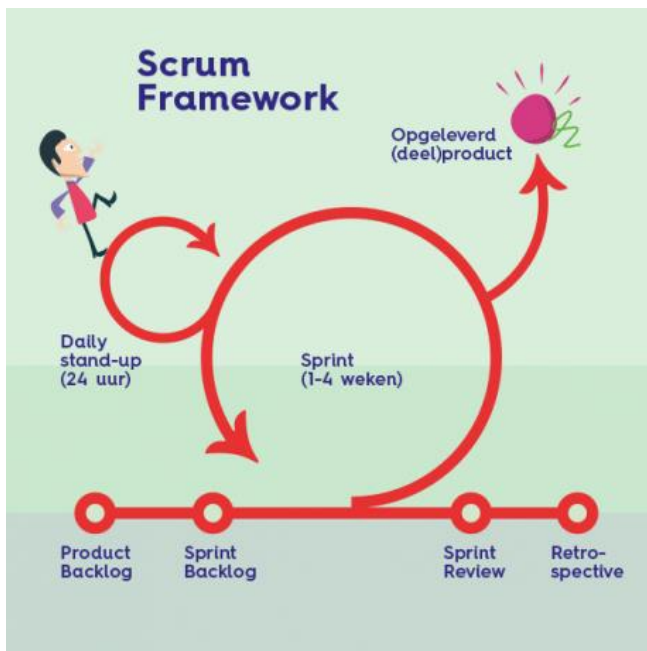
Het SCRUM-team bestaat over het algemeen uit 3-9 leden. Voor dit project bestaat het SCRUM-team uit 8 teamleden. Het SCRUM-team bestaat uit de volgende rollen:

- **SCRUM-Master:** De SCRUM-Master begeleidt het team door ervoor te zorgen dat het proces op de juiste manier wordt gevolgd. Hij/Zij is verantwoordelijk voor de vergaderingen en het regelen van voorzieningen zoals werkruimte, software en hardware. Als de SCRUM-Master zijn werk goed uitvoert, kunnen de teamleden ongestoord hun werk doen.
- **Product Owner:** De Product Owner is de opdrachtgever, hij heeft er het grootste belang bij dat het product gemaakt wordt en van goede kwaliteit is. De Product Owner beheert de Backlog en bepaald hiermee wat in welke volgorde wordt gedaan.
- **Teamleden:** Het team is zelf organiserend. Het team bepaald zelf hoe de items uitgevoerd moeten worden. Het team is cross-functioneel, dit betekent dat de leden van het team samen minimaal 80% van de vaardigheden bezitten om de taken te kunnen uitvoeren. Het team is samen verantwoordelijk voor het resultaat, ook al worden de taken uitgevoerd door teamleden die de benodigde vaardigheden bezitten.
- **Stakeholders:** Dit zijn de belanghebbende van het project. De stakeholders krijgen resultaten te zien en geven feedback op de producten.

Sprints/Iteraties

Met de SCRUM-methode wordt dit project verdeelt in vier sprints van twee weken. Figuur 3 weergeeft hoe elke sprint doorlopen wordt. De Sprint Retrospective is tijdens dit project geen formele bijeenkomst, maar wordt informeel besproken tussen de teamleden. Elke sprint bestaat uit drie onderdelen:

- **Sprint planning:** Tijdens de sprint planning wordt het werk dat gedaan moet worden ingepland. Hier is het hele team bij aanwezig. De Scrum Master heeft de verantwoordelijkheid dat alles duidelijk is voor ieder Scrum lid. In de meeting wordt besproken wat er gedaan gaat worden tijdens de aankomende sprint. Dit wordt aan de hand van de Product Backlog vastgelegd. Nadat is bepaald wat het doel is van de aankomende sprint, bepaald het projectteam hoe elk item uit de Product Backlog uitgevoerd gaat worden.
- **Daily stand-up:** De daily stand-up is een bijeenkomst die dagelijks of elke X aantal dagen per week wordt gehouden. Deze daily stand-up duurt maximaal. Het doel van deze bijeenkomst is om de werkzaamheden op elkaar af te stemmen voor de komende dag(en). Tijdens de daily stand-up worden de volgende drie vragen beantwoord door ieder teamlid:
 1. Wat heb ik bereikt sinds de vorige daily stand-up?
 2. Wat ga ik vandaag bereiken?
 3. Verwacht ik obstakels en kan een teamlid mij hierbij helpen?
- **Sprint review:** De sprint review wordt aan het einde van een sprint gehouden. Deze meeting is om de gemaakte resultaten te presenteren aan de Product Owner. De sprint review duurt maximaal 4 uur en het hele Scrum team en de belanghebbende zijn aanwezig. Het Scrum team presenteert wat er gerealiseerd is tijdens de sprint en de Product Owner verteld welke Product Backlog items voltooid zijn en welke nog niet af zijn. De belanghebbende hebben hierbij de kans om feedback te geven.



Figuur 3 scrum sprint framework

3.3 MoSCoW

De MoSCoW methode is tijdens de opdracht gebruikt om de wensen en eisen van de opdrachtgever, key-users en stakeholders te verzamelen en prioriteren. Deze methode wordt veel gebruikt binnen de productontwikkeling, softwareontwikkeling en bedrijfsanalyses. De eisen van een project worden in de volgende categorieën ingedeeld:

1. **Must-have:** Vereist om te kunnen spreken van een werkbaar product;
2. **Should-have:** Hoge prioriteit, maar niet vereist voor een bruikbaar product;
3. **Could-have:** Optie die alleen wordt meegenomen als er tijd over is;
4. **Would-have:** Geen prioriteit, kan eventueel in de toekomst opnieuw worden overwogen

De MoSCoW methode is een onderdeel geweest in het achterhalen en het prioriteren van de requirements voor de opdracht. De reden dat er voor de MoSCoW methode is gekozen, is omdat de stakeholders veel verschillende belangen hebben bij de opdracht. Door de MoSCoW methode toe te passen worden enkel de meest belangrijke eisen toegepast. De wensen/eisen die minder belangrijk zijn die kunnen later worden ontwikkeld, maar vallen niet binnen de scope van deze opdracht. De resultaten zijn in de DevOps tool verwerkt.

3.4 DAMA DMBOK2 Raamwerk

Volgens de organisatie DAMA kan het DMBOK2 raamwerk toegepast worden om datakwaliteit management binnen een organisatie te optimaliseren. Als het raamwerk op de juiste wijze wordt toegepast binnen de organisatie zijn volgens DAMA alle facetten van datakwaliteit management gedekt.

Meer informatie over het DAMA DMBOK2 meetinstrument is te vinden in het document “meetinstrument datamanagement”. In dit document zijn alle domeinen uitgebreid beschreven die betrekking hebben en relevant zijn voor de organisatie.

Tijdens dit project worden niet alle onderdelen van het DAMA DMBOK2 raamwerk uitgevoerd. Alle domeinen worden geïntroduceerd en beschreven, met de meerwaarde van elk domein voor de organisatie. Vervolgens wordt er ingezoomd op een aantal domeinen die van groot belang zijn om die op orde te hebben op korte termijn. Tenslotte wordt er nog verder ingezoomd op de datakwaliteit en wordt hier een proof of concept op ontwikkeld. Dit proof of concept is ontwikkeld op basis van het DAMA DMBOK2 raamwerk.

3.5 Microsoft DevOps

DevOps is een online tool om mensen, processen en technologieën samen te brengen. DevOps, een samenstelling van dev (development) en ops (operations), is een softwareontwikkelingspraktijk waarbij de ontwikkeling en IT-activiteiten worden gekoppeld. DevOps wordt tijdens het project gebruikt om met de SCRUM-methode te werken.

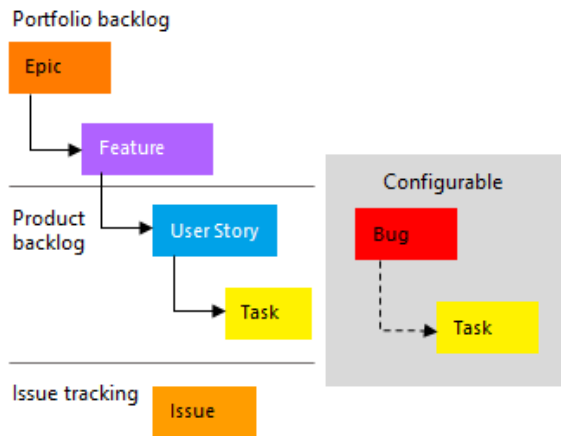
DevOps bevat het plannen, bijhouden, ontwikkelen, bouwen en testen. Tijdens het voorbereiden voor de opdracht, zijn alle requirements en activiteiten in DevOps geplaatst. DevOps is met samenwerking tussen de stagebegeleider en de afstudeerders opgezet. De opzet is als volgt:

Begrip	Omschrijving
Overview	Is het dashboard of een samenvatting van de huidige persoonlijke en gezamenlijke voortgang. Dit dashboard is gemaakt om te kunnen zien welke activiteiten/issues nog open staan en welke zijn afgerond.
Product Backlog	In de product backlog staan alle activiteiten, taken, user stories, sprints en issues van de opdracht. Deze zijn ingedeeld per onderdeel en hebben allen een prioriteit.
Work items	Hieronder vallen alle activiteiten, user-stories, taken en bugs van de opdracht (Zie Figuur 4 work items DevOps)

Tabel 5 DevOps begrippen

De work-items zijn zoals Figuur 4 geordend. Voor de opdracht zijn er definities gemaakt voor de work-items:

- **Epic:** De Epic is het hoogste niveau van project gerelateerde onderdelen.
- **Features:** Is een lager niveau, waarbij context van de onderliggende activiteiten of eisen duidelijk worden.
- **User-stories:** In de user stories staan de wensen en de eisen van de stakeholders. De user stories zijn opgesteld volgens de INVEST-checklist:
 - **I - Is it Independent:** Zorg dat de user stories onafhankelijk zijn van andere stories, zo kan de product owner prioriteiten stellen.
 - **N - Is it Negotiable:** User stories moeten niet te gedetailleerd zijn. Ontwikkelaars moeten de vrijheid krijgen om zelf in te vullen hoe een functionele wens vorm krijgt.
 - **V - Is it Valuable:** De functionaliteit moet uiteindelijk een waardevolle bijdrage leveren voor de gebruiker.
 - **E - Is it Estimable:** Kun je inschatten hoe veel tijd en werk het ongeveer kost om de functionaliteit te ontwikkelen en is alle kennis in huis?
 - **S - Is it Small enough:** Beschrijft de user story maar een functionaliteit? Het moet uiteindelijk ook niet te klein zijn dat het te weinig waarde zal opleveren.
 - **T - Is it Testable:** Is de acceptatiecriteria omschreven bij de user story, waardoor je kunt bepalen of de functionaliteit straks aan de eisen voldoet?
- **Tasks:** Dit zijn taken die worden toegewezen aan meestal één individu. Meerdere taken kunnen een user story beantwoorden, of bepaalde taken zijn nodig om bijvoorbeeld bepaalde rollen toegang te geven tot systemen of data om taken uit te kunnen voeren.
- **Bugs:** Zijn problemen of errors die gedurende de opdracht kunnen opduiken. Het oplossen van bugs wordt beschreven in één of meerdere tasks.



Figuur 4 work items DevOps

4 Onderzoeksvragen

De hoofdvraag wordt beantwoord aan de hand van de deelvragen. Wanneer alle deelvragen zijn beantwoord heeft de lezer voldoende kennis om de hoofdvraag te kunnen beantwoorden.

Het onderzoek is in opdracht vanuit de inkoopafdeling, met verwijzingen naar “de organisatie” en “onderneming” wordt de inkoopafdeling bedoelt. Deelvraag 6 is op een andere manier uitgewerkt dan de andere deelvragen. Deze deelvraag doorloopt meerdere activiteiten om tot een eindresultaat te komen.

Hoofdvraag

Welke maatregelen dienen gemaakt te worden om de datakwaliteit te waarborgen wanneer de data toegankelijkheid verhoogd wordt.

Deelvragen

1. Welke KPI's en requirements zijn gedefinieerd binnen de inkoopafdeling?
2. Wie zijn de stakeholders en welk belang hebben zij bij het project?
3. Wat doet de organisatie aan datamanagement?
4. Welke invloed heeft de data migratie naar Microsoft Azure op het project?
5. Welke aspecten van datamanagement hebben een impact op dit project?
6. Hoe wordt datakwaliteit management getoetst?
7. Welke tools kunnen gebruikt worden voor datakwaliteit management?
8. Hoe worden data issues verbeterd?
9. Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?

5 Onderzoekresultaten

De onderzoekresultaten worden per deelvraag beantwoord. Na het beantwoorden van de deelvragen wordt er een conclusie getrokken en wordt de hoofdvraag beantwoord.

De individuele deelvraag zijn in de meeste gevallen verdeeld in twee paragrafen. Het eerste deel aanpak, beschrijft op welke manier de deelvraag is uitgevoerd. Vervolgens worden de resultaten beschreven. In het hoofdstuk “B-Competenties” staat beschreven welke B-competenties zijn aangetoond bij het beantwoorden van iedere deelvraag.

5.1.1 Welke KPI's en requirements zijn gedefinieerd binnen de inkoopafdeling?

Croonwolver&dros werkt met KPI's om bij te houden hoe de organisatie en haar bedrijfsprocessen presteert. Doelen worden gedefinieerd en op basis van de doelstellingen worden KPI's vastgesteld. De KPI's krijgen een target vanuit de strategische top. Op basis van de KPI's en de targets worden de status en verbeteringen of verslechtingen in het bedrijfsproces aangetoond.

Het doel van het opstellen van KPI's en requirements voor dit project, is om informatie te identificeren en te prioriteren. Hoe frequenter een veld voorkomt in de KPI's en requirements, hoe belangrijker dit veld zal zijn voor de rapportages.

Aanpak

De KPI's en requirements zijn doormiddel van een workshop, interviews en meetings met stakeholders gedefinieerd en geprioriteerd. De requirements zijn gevalideerd tijdens de eerste sprint review met de stakeholders en de opdrachtgever (Zie [Bijlage 1 Requirements en KPI's in DevOps](#)).

Elke KPI en requirement heeft een prioriteit meegekregen conform de MoSCoW methode. Dit is tijdens de workshop met de stakeholders vastgelegd ([MoSCoW](#)).

De KPI's krijgen een indexnummer, een prioriteitsniveau en zijn ingedeeld in verschillende aandachtspunten. De aandachtspunten zijn kostenbesparing, kwaliteit en management. Iedere KPI toont een relevante waarde aan waar het desbetreffende aandachtspunt actie op kan uitvoeren. De target per KPI is later te bepalen door de directie. De KPI's moeten per maand en per kwartaal de waarden kunnen weergeven.

De requirements die zijn geformuleerd zijn user requirements, om de requirements in een later stadium terug te kunnen vinden krijgen ze een indexnummer mee. Vervolgens staat er van welke functie de requirement komt. De structuur van de user requirements zijn als volgt:

“Als <gebruiker/rol> wil ik <functie/actie>, zodat ik <doel/waarde>.”

De KPI's zijn in drie verschillende aandachtspunten gecategoriseerd. Dit is gedaan om onderscheid te maken tussen de focusgebieden waarde de KPI op stuurt:

- **Kostenbesparing:** Deze KPI's richten zich op voornamelijk op contracten en bonussen. Wanneer een bonustarget wordt behaald zijn er afspraken om een X-aantal percentage korting te ontvangen bij een leverancier. Bestellen via contracten is goedkoper dan bestellen bij leveranciers zonder contracten.
- **Kwaliteit:** Richt zich vooral op de kwaliteit van de leveranciers. Hierbij komen vraagstukken zoals: Levert leverancier X op tijd, klopt de bestelorder met de factuur, is de levering compleet of hoe is de algemene ervaring met de leverancier. Leveranciers die een slechte leverperformance of kwaliteit leveren, kost de organisatie tijd en geld.
- **Management:** Gericht op het analyseren van bedrijfsprocessen die impact hebben op de omzet. De doelstellingen worden bepaald door de business professionals en worden bijgehouden door het management. (Leansixsigma, Leansixsigma, 2019)

Resultaat

De geformuleerde KPI's zijn ontwikkeld voor de inkoopafdeling. De KPI's zijn op strategisch niveau geformuleerd, om de performance van de inkopers te analyseren. Het doel is dat de inkoopafdeling de bedrijfsproces-prestaties kan aflezen van het dashboard en beslissingen kan maken op basis van data. In Tabel 6 zijn de KPI's van de inkoopafdeling weergegeven.

Prioriteit	Aandachtspunt	Key Performance Indicator Inkoop
M	Kostenbesparing	Percentage van het inkopen via raamcontracten (Aantal + Bedragen)
M	Kostenbesparing	De huidige status van inkopen bij leveranciers vs. De bonus targets
S	Kwaliteit	Percentage leverperformance per leverancier (gewenste leverdatum vs. daadwerkelijke leverdatum)
S	Kwaliteit	Percentage productkwaliteit score per leverancier (volledig/niet volledig)
M	Management	Aantal openstaande facturen per leverancier
W	Management	Omzet percentage van een VOF (waar Croonwolter&dros penvoerder van is)
M	Management	Omzet percentage van inkopen via voorkeurleverancier vs. Niet-voorkeurleverancier
W	Management	Percentage/Aantal Status werkbegroting (Besteden VS. nog te besteden)
C	Kostenbesparing	Omzet Percentage met bestelorder
S	Management	Percentage aantal leveranciers die leveren (VS. Totale leveranciers)

Tabel 6 Key performance indicatoren

De requirements zijn vervolgens in Microsoft DevOps geplaatst, van waaruit de planning en activiteiten beheerd worden. Zie [Bijlage 1](#) voor de complete requirement en KPI lijsten.

5.1.2 Wie zijn de stakeholders en welk belang hebben zij bij het project?

Het doel van deze deelvraag is om te achterhalen wie de stakeholders zijn en hoe betrokken zij zijn bij het project.

Aanpak

Bij het realiseren van de opdracht zijn er stakeholders betrokken. De stakeholders zijn gedefinieerd in interne/externe stakeholders ([Leansixsigma](#), [Leansixsigma](#), 2019) en primaire/secundaire stakeholders. De stakeholdersanalyse is gevalideerd door de opdrachtgever en de stagebegeleider.

- De interne stakeholders staan dicht bij het project. Zij zijn direct betrokken en werken mee aan het eindresultaat.
- De externe stakeholders werken niet mee aan het project, maar zijn er wel bij betrokken en hebben belang bij het project.
- Primaire stakeholders zijn van redelijk tot groot belang voor het project.
- Secundaire stakeholders hebben geen direct belang bij het project, maar zijn wel van invloed op de uitvoering van het project.
- De belangen van de stakeholders worden beschreven in de “Belang & Impact” kolom. De impact wordt aangegeven met een cijfer van 1 t/m 5, waarbij stakeholders met een 1 het minst impact & belang hebben op het project en 5 het grootste belang & impact hebben.

Resultaat

De stakeholderanalyse is gevalideerd door de Business Intelligence Manager, wie tevens ook een project lid en stagebegeleider is.

Stakeholder	Intern/Extern	Primair	Functie	Belang & Impact
Mathieu Koomans v.d. Dries	Intern	Primair	Opdrachtgever	5
Susan Verbeij	Intern	Primair	Consultant, Tester	5
Rob Kabel	Intern	Primair	Projectbegeleider/Tester	4
Jeroen Sman	Extern	Secundair	Informatie manager ICT	2
Managementteam	Extern	Secundair	Stuurgroep	2
Albert Huttenga	Intern	Primair	Informatieanalist	4
TBI Holding	Extern	Secundair	Overkoepelende onderneming van CW&D	2
Eltjo Renken	Intern	Primair	Databasebeheerder	4
Sike v. Dongen	Intern	Primair	Afstudeerder	4

Tabel 7 stakeholderanalyse

De primaire en secundaire stakeholders hebben een belang bij het project Het management van inkoop heeft tijdens een presentatie ([Bijlage 10 management presentatie](#)) en een demo ([Demo Croonwolver&dros inkoopdashboard](#)) van het inkoopdashboard feedback gegeven. Het management team is enthousiast dat er een dashboard raamwerk wordt ontwikkeld die over de breedte van de organisatie geïmplementeerd gaat worden, waarmee ieder dashboard van Croonwolver&dros dezelfde “look and feel” heeft, maar realiseert zich ook dat de dashboards geen waarde hebben als de datakwaliteit niet op orde is.



Datakwaliteit speelt een grotere rol bij de strategische top dan op operationeel niveau. Het management kijkt met een bepaalde visie naar de toekomst en erkent het belang van datakwaliteit. De stuurgroep ziet het als noodzaak dat er een datakwaliteit management tool komt. Het belang van de stuurgroep is dat zij strategische beslissingen willen kunnen maken op basis van data, iets wat op dit moment niet vanzelfsprekend is. Het valideren van gegevens en analyses kost op dit moment meer tijd en energie dan gewenst.

5.1.3 Wat doet de organisatie aan data management?

Aanpak & Resultaat

Croonwolver&dros is na de fusie significant meer data gaan genereren. De data wordt beheerd in het ERP-systeem van Croonwolver&dros. Een organisatie als Croonwolver&dros heeft meerdere administraties en verwerkt duizenden regels per dag in het ERP-systeem. Als hier geen databeheer voor is zal de organisatie op ten duur in de problemen raken. Om meer informatie over dit onderwerp te verzamelen, is er een interview afgenomen van een inkoper die de data beheert, zonder een formele titel hiervoor te hebben. Hierover werd het volgende gezegd door de data steward van de inkoopafdeling:

Foutieve data wordt handmatig ontdekt. Uit de analyse is gebleken dat er geen gestructureerde manier is om foutieve data op te lossen. De foutieve data wordt gevonden door analyses uit te voeren op de data en vervolgens een afwijkend resultaat terug te krijgen ten opzichte van de werkelijkheid. Het is aan de persoon zelf of hij/zij een melding maakt van de afwijkende data. De desbetreffende databeheerder bekijkt de melding en heeft vervolgens twee opties:

- 1. De foutieve data is op te lossen door dit handmatig in Acto te wijzigen. Er wordt een melding gemaakt bij één van de drie beheerders, die vervolgens de foutieve data wijzigt, dit wordt op veldniveau gedaan.*
- 2. De data correctie gaat te veel tijd kosten, de databeheerder neemt contact op met Acto. Acto is de leverancier van het ERP-systeem. De verantwoordelijkheid hiervan ligt bij Acto, om ervoor te zorgen dat dit niet meer voorkomt in de toekomst. Deze maatregel wordt in de praktijk zelden toegepast.*

5.1.4 Welke invloed heeft de data migratie naar Microsoft Azure op het project?

Croonwolver&dros is vanaf begin 2019 begonnen met de migratie van het datawarehouse On-premise naar Microsoft Azure. De migratie wordt uitgevoerd door de database administrator en de data-analist. De migratie van On-premise naar de Cloud gaat maanden duren. Het doel van onderzoeken wat de invloed is van de migratie naar Azure is om risico's te identificeren, maar ook om kansen te benutten op het gebied van services en governance met het overstappen naar Azure.

Aanpak

Het onderzoek naar de invloed van de data migratie naar Azure is uitgevoerd volgens een kwalitatief onderzoek uitgevoerd. In overleg met de data-analist, de databasebeheerder en een consultant van Motions10, zijn de volgende resultaten voortgekomen.

Resultaat

Het huidige BI landschap staat weergegeven in [Figuur 5 ETL proces on-premise](#). Het huidige BI landschap bevat problemen die niet eenvoudig op te lossen zijn in de huidige situatie. De problemen zijn in verschillende categorieën ingedeeld:

Security

Toegangsrechten zijn per gebruiker per database ingesteld, hierdoor is kost het toevoegen van nieuwe gebruikers veel tijd. Een overzicht geven van gebruikers die toegang hebben tot een bepaald object is erg arbeidsintensief. Het huidige securitymodel maakt het moeilijk om inzicht te krijgen in het gebruik van de data.

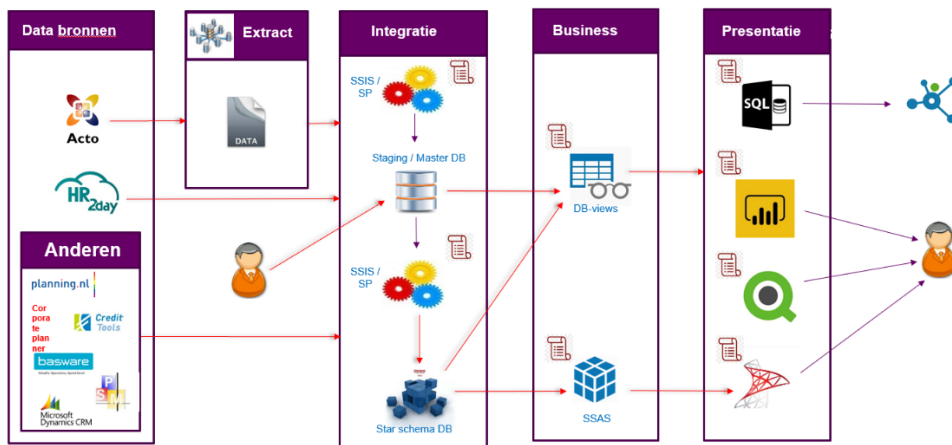
Operationeel

Het huidige datawarehouse zit tegen zijn plafond aan, waardoor er performance- en stabiliteit problemen voorkomen. Het ophalen van de data elke nacht loopt uit zijn tijdslot. Sommige rapportages worden niet automatisch verversd waardoor dit handmatig gedaan moet worden. Er is geen testomgeving waardoor testen in de productieomgeving plaats vindt. Dit leidt tot problemen voor de gebruikers. De BI-afdeling heeft niet de middelen om alle processen te bewaken in de huidige situatie.

Architectuur

In de loop van de tijd zijn er verschillende componenten toegevoegd aan het landschap zoals Power BI, staging area, SSAS en het datawarehouse. Dit is gedaan om snel nieuwe features te realiseren, zodat de business door kan blijven ontwikkelen. Op een gegeven moment wordt de architectuur een moeras waarin je de weg niet meer weet terug te vinden. Deze problemen maken het moeilijk om Data Governance toe te passen.

BI LANDSCHAP PER OKTOBER 2018



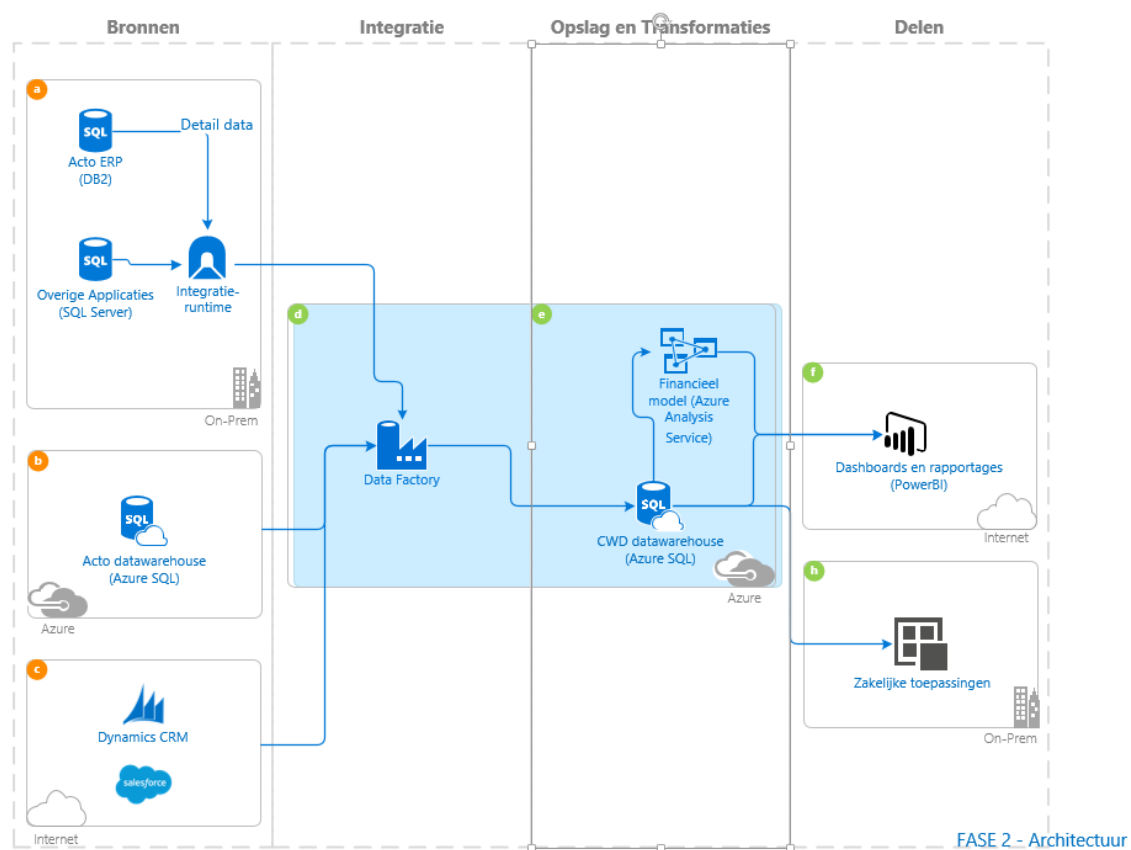
Figuur 5 ETL proces on-premise

In Figuur 6 is weergegeven hoe het toekomstige ETL proces er uit gaat zien. Het grote verschil tussen figuur 5 en figuur 6 is dat er in figuur 6 significant minder software gebruikt is. Het achterhalen van problemen zal makkelijker worden en de snelheid van het verwerken van data is volledig schaalbaar naar de wensen van de organisatie.

Het Acto ERP-systeem levert 80% van alle data van Croonwouter&dros. Het Financieel model is een nog onzekere service tussen PowerBI en het datawarehouse speciaal voor de Finance afdeling om gemakkelijk rapportages te kunnen bouwen op data uit het CWD DWH.

Verklaring van de letters in Figuur 6 ETL-proces Azure:

- A. De on-premises bronnen zoals het Acto ERP en Basware.
- B. Het Acto datawarehouse (DWH) gehost in Azure als een software as a service (SAAS) dienst.
- C. Externe webapplicaties zoals bijvoorbeeld Dynamics CRM of HR2day.
- D. Na kantooruren zal Data Factory (ADF) de verschillende bronnen (A, B, C) behalve het Acto ERP uitlezen en het CWD datawarehouse bijwerken met de laatste stand van zaken.
- E. Het Acto ERP wordt in Azure gerepliceerd om te kunnen voorzien in near real-time (NRT) data behoefte.
Het CWD-datawarehouse is een Azure SQL-database waar alle informatie staat opgeslagen. De informatie wordt getransformeerd d.m.v. SQL stored procedures en aangeboden d.m.v. views. Het Financieel model is een nog onzekere service tussen PowerBI en het datawarehouse speciaal voor de Finance afdeling om gemakkelijk rapportages te kunnen bouwen op data uit het CWD DWH.
- F. PowerBI kan data uitlezen uit het CWD DWH en het financieel model in Azure Analysis Services en staat gebruikers in staat om interactieve rapportages te bouwen.
- G. Zakelijke toepassingen zoals Active Directory (AD) en Topdesk kunnen speciale views in het CWD datawarehouse uitlezen voor integratie doeleinden.



Figuur 6 ETL-proces Azure

De migratie naar een Cloud omgeving heeft voordelen, die zijn als volgt:

- **Flexibiliteit:** Het datawarehouse kan op ieder moment worden uitgebreid of gereduceerd worden naar de gewenste grootte. Ook kunnen er meer of minder processoren worden ingesteld om de data te verwerken naar wensen van de gebruiker. Servers kunnen worden aangeschaft of afgeschaft zonder dat je de server hoeft te kopen of verkopen.
- **Onderhoudskosten:** Met het gebruik van een Cloud omgeving heeft de organisatie geen onderhoudskosten van servers. Bij Azure betaal je Pay-per-Use: Je betaalt waar wat je gebruikt. Dit houdt in dat er geen grote investeringen nodig zijn om een datawarehouse in te richten.
- **Betrouwbaarheid:** Een Cloud service garandeert dat de server 99.9% van de tijd uptime is, waar er On-premise geen garantie is.
- **Security:** Het Azure platform heeft veel certificeringen op het gebied van Security. De gebruiker is zelf eindverantwoordelijk voor de beveiliging van zijn systemen, maar Microsoft heeft alle mogelijkheden om je omgeving te beveiligen. Role Based Access Control (RBAC) binnen Azure draagt bij aan het beheersbaar maken van de toegangsrechten.
- **Verbinding:** Met Azure is er geen virtual machine/VPN verbinding nodig om toegang te krijgen tot bepaalde data. De data is bereikbaar via het publieke internet.
- **Services:** Microsoft Azure heeft meer dan honderd services voor alle computersystemen en in alle programmeertalen. Hiermee kun je snel services toevoegen of verwijderen naar de behoeften van de organisatie.
- **Architectuur:** De services bevinden zich in één platform. Dit draagt bij aan het toepassen van Data Governance in de organisatie. Het herleiden van problemen hoeft niet meer via alle losse componenten, maar gebeurt nu binnen één platform.

5.1.5 Welke aspecten van datamanagement hebben een impact op dit project?

Aanpak

Met het beantwoorden van deze vraag wordt er ingegaan op de drie domeinen die relevant zijn aan dit project, maar ook over de gehele organisatie spelen. In het theoretisch kader wordt beschreven welke aspecten van de drie datamanagement domeinen volgens het DAMA DMBOK2 belangrijk zijn voor de organisatie. Over de drie domeinen wordt beschreven welke maatregelen er genomen dienen te worden om datamanagement binnen de organisatie te verbeteren. Er wordt een stap dieper ingegaan op het domein Datakwaliteit management, omdat dit domein centraal staat in het project en hier een product van wordt geïmplementeerd.

Figuur 7 geeft de relatie tussen de drie domeinen weer. Data Governance is een overkoepelend domein die raakvlakken heeft met de andere data management domeinen.



Figuur 7 Data Management drie domeinen

Resultaat

Data Governance

Data Governance is ervoor te zorgen dat gegevens correct worden beheerd, volgens beleid en best practices. Data Governance focust zich op de manier waarop beslissingen worden genomen over gegevens en hoe mensen en processen met de gegevens omgaan. Er is op dit moment geen Data Governance management binnen Croonwolter&dros.

Data is een waardevol bezit, om deze waarde te behouden of te verhogen moet dit onderhouden en aangepast worden naar de wensen en eisen van de organisatie. Waarom is Data Governance dan belangrijk voor Croonwolter&dros? Croonwolter&dros bestaat uit vijf verschillende divisies. De werknemers van de divisies bevinden zich op verschillende locaties over heel Nederland. Croonwolter&dros heeft verschillende administraties die onder de organisatie werken. Met al deze

variabelen zijn er veel verschillende waarheden, definities, procedures en werkwijzen. Dit zorgt voor problemen bij het samenvoegen van de gegevens.

Een vereist onderdeel van Data Governance is één of meerdere data eigenaren in te stellen. Zonder een data eigenaar zal niemand zich verantwoordelijk of geroepen voelen om de waarde van de data te onderhouden. Om definities, procedures, strategie, beleid, normen & kwaliteit en naleving over de gehele organisatie uit te voeren, dient Croonwolter&dros een Data Governance council in te stellen. Deze council is samengesteld uit vertegenwoordigers vanuit alle business units. Het Data Governance council stelt het databeleid vast en de procedures en normen voor de organisatie als één geheel. Dit moet gelijk zijn over de gehele organisatie.

Het volgende element is om een effectief databeleid en procedures integraal en over de gehele organisatie beschikbaar stelt. Het is van belang om een gemeenschappelijk informatiemodel te hebben voor de hele organisatie. Anders blijft iedereen zijn eigen werkwijze handteren.

Binnen de organisatie is het fundamenteel om definities vast te stellen. Definities in verschillende vormen:

- Meta-data management
- Wie is waar verantwoordelijk voor
- Wie doet welke werkzaamheden
- Procedures, beleid, (kwaliteit)normen en strategie dienen voor iedereen binnen de organisatie beschikbaar te zijn
- Invoerprocedures voor data beheerders
- Een Data issue procedure, dit is van belang zodat na correcties zijn uitgevoerd op data issues de gegevens nog steeds overeenkomt met de oorspronkelijke originele data elementen

Tenslotte dienen de doorgevoerde verbeteringen nageleefd te worden. Dit geldt niet enkel voor de werkwijze of procedures, maar ook voor de wet en regelgeving (AVG) en om de waarde van de data te waarborgen. De datakwaliteit speelt hierin een centrale rol.

Datakwaliteit management

De kwaliteit van de data zal niet direct een impact hebben op de business. De organisatie heeft jarenlang ogenschijnlijk zonder datakwaliteit management succesvol kunnen opereren. Waarom nu starten met datakwaliteit management? Met de continue groei van de organisatie, groter wordende beschikbaarheid van data en de vraag van medewerkers naar betrouwbare data om sneller beslissingen te nemen wordt data langzaam maar zeker een “asset” die aandacht moet krijgen. Wanneer de data volume groeit met het aantal databronnen waaruit de data wordt geladen, gaat de datakwaliteit achteruit. Dit probleem kan voorkomen worden met een datakwaliteit management service. Het doel van dit datakwaliteit management product voor Croonwolter&dros is:

- Gebruikers bewust te maken over het belang van datakwaliteit
- Definiëren van data rules en specificaties voor datakwaliteit controles
- Het implementeren van een datakwaliteit tool waarmee foutieve geïdentificeerd en opgelost kunnen worden door gebruikers.
- Definiëren en implementeren van processen voor het meten, waarborgen en rapporteren van het datakwaliteitsniveau
- Identificeren van fouten in de data die handmatig (in de meeste gevallen) niet gedetecteerd zullen worden

De tool wordt geïmplementeerd in de inkoopafdeling. De databasebeheerders van de inkoopafdeling kunnen zelfstandig de foutieve data met behulp van deze tool oplossen. De tool is duurzaam ontworpen, wat betekent dat het schaalbaar is en in een later stadium de rest van de organisatie geïmplementeerd kan worden.

Het hoofdstuk “Hoe wordt de datakwaliteit getoetst?” Gaat dieper in op hoe de datakwaliteit tool is ontworpen, ontwikkeld en geïmplementeerd.

Data Security

Op basis van het theoretisch kader is er een antwoord gegeven op het antwoord welke relevante Data Security aspecten raakvlakken heeft op het project. Meer informatie over Data Security is terug te vinden in het theoretisch kader.

Er is geen voorgeschreven manier om Data Security te implementeren in een organisatie. Organisaties moeten hun eigen beveiligingsmaatregelen ontwerpen en aantonen dat de organisatie minimaal voldoet aan de vereisten van de wet- of regelgeving. Croonwolver&dros is actief bezig om Data Security te implementeren. Microsoft Azure heeft services om Data Security te implementeren en onderhouden.

Croonwolver&dros heeft adequate maatregelen genomen om de data te beveiligen tegen buitenaf. Gevoelige data (High Risk Data) is alleen bereikbaar vanaf het netwerk binnen de organisatie, via een VPN verbinding met authenticatiebeveiliging of via Citrix. Als iemand een account heeft in de organisatie en toegang heeft tot het netwerk, heeft die persoon toegang tot de netwerkschijf. Binnen de netwerkschijf heeft deze persoon enkel toegang tot de bestanden waar toegang tot is gegeven. Er is een Role Based Access Control management aanwezig voor bepaalde rapportages. Dit houdt in dat werknemers een rol of functie hebben, waarmee mensen met deze rol/functie bepaalde documenten, data of systemen wel of niet mogen lezen en/of bewerken. Dit wordt niet per persoon bepaald maar per rol of functie.

Binnen het datawarehouse is geen Role Based Access Control management aanwezig. Een business medewerker krijgt geen toegang tot het datawarehouse. Om toegang te krijgen tot het datawarehouse, dienen de betreffende personen te beschikken over een VOG. Als iemand toegang heeft tot het datawarehouse kan deze persoon alle tabellen bekijken, met uitzondering van data die onder de AVG-wetgeving valt. Er zijn wel restricties opgelegd dat de tabellen niet gewijzigd kunnen worden.

Met het ingaan van het Algemene Verordening Gegevensbescherming (AVG) in mei 2018 is Croonwolver&dros verplicht om de richtlijnen van de overheid aan te houden. Er is een externe partij ingehuurd om ervoor te zorgen dat Croonwolver&dros zich aan de wettelijke eisen houdt omtrent persoonsgegevens. Deze partij onderzoekt de huidige status van Croonwolver&dros en heeft als doelstelling om minimaal aan alle wettelijke verplichtingen van het AVG aan het einde van het jaar te voldoen.

Data management is een nieuw begrip binnen de organisatie. De bewustwording van het “managen van data” is nog niet bij iedereen in de organisatie geland. Om in de toekomst de efficiëntie omtrent het gebruik en managen van data binnen de organisatie te verhogen, is het noodzakelijk om een Data Governance project te starten. In dit project is de doelstelling om over de hele organisatie duidelijk te

maken wie wat doet en waar hij/zij verantwoordelijk voor is. De manier van het omgaan met data dient te worden gedefinieerd doormiddel van mensen begeleiden en beleid & procedures opstellen. In dit project dient voor de datakwaliteit een tool/product aangeschaft of ontwikkeld te worden om dit te kunnen waarborgen. Met het ingaan van de AVG en de verhoogde risico voor digitale inbrekers “hackers” dient de organisatie zich minimaal aan de wettelijke vereisten te houden.

5.1.6 Hoe wordt de datakwaliteit getoetst?

De toetsing van de datakwaliteit is op basis van het DAMA DMBOK2 raamwerk ontworpen en uitgevoerd. Het raamwerk biedt een functioneel kader voor het ontwerpen van de tool op een gestructureerde manier. Het functioneel kader is samengevoegd met de Business waarmee Croonwolter&dros een nieuw bedrijfsproces heeft geïmplementeerd. Doormiddel van het doorlopen van zes activiteiten wordt deze datakwaliteit management tool geïmplementeerd.

5.1.6.1 Definieer data van hoge kwaliteit

Datakwaliteit management is een bedrijfsproces dat structureel uitgevoerd dient te worden. Voordat er data gemeten wordt of normen worden opgesteld, dienen er definities van datakwaliteit gedefinieerd te worden. Onderstaande vragen identificeren pijnpunten, bedrijfsbehoeften, termen en prioriteiten voor de verbetering van datakwaliteit. De antwoorden op de vragen betreffen de interne visie bij Croonwolter&dros.

1. Wat wordt er bedoelt met “hoge datakwaliteit”?

Met hoge datakwaliteit wordt bedoeld dat de data betrouwbaar genoeg is om strategische inkoopbeslissingen te nemen. De eisen die gesteld worden verschillen per data element (attribuut). Croonwolter&dros heeft nog geen concrete metrics bepaald per data element of per dataset. Daarom geldt op dit moment “zo hoog mogelijk”.

2. Wat is de impact van lage kwalitatieve data op bedrijfsactiviteiten en strategie?

Croonwolter&dros wordt steeds meer afhankelijk van data. Dit geldt ook zeker voor de afdeling Inkoop.

Lage kwaliteit data kan leiden tot foutieve beslissingen bij:

- Het plaatsen van bestellingen
- Het maken van kortingsafspraken met leveranciers
- Het beoordelen van inkoopgedrag door inkopers en projectleiders
- etc.

3. Hoe kan data van hoge kwaliteit bijdragen aan de inkoopstrategie?

Door het beschikbaar hebben van hoge kwaliteit data kunnen de onderhandelingen met leveranciers over b.v. leveringsvoorwaarden effectiever plaatsvinden.

Door inzicht te hebben in het product-/dienstenaanbod en betrouwbare historische data kan de gekozen strategie waarbij zoveel mogelijk met een relatief kleine groep voorkeurleveranciers wordt gewerkt, gerealiseerd worden.

4. Welke prioriteiten heeft het verhogen van de datakwaliteit?

De prioriteit is hoog maar de realiteit is dat er nog niet naar gehandeld wordt als gevolg van beschikbaarheid van resources.

Het is aan de data-eigenaar om toe te zien dat structurelere toetsing van de datakwaliteit onderdeel is van het normale bedrijfsproces.

Tools kunnen hierbij ondersteuning bieden. De verwachting is dat de Inkoopafdeling wel de benodigde prioriteit zal geven.

5. Welke knelpunten ondervindt de (inkoop)afdeling op dit moment?

Gerelateerd aan data, ondervindt de Inkoopafdeling momenteel een groot gemis aan betrouwbare rapportages die inzicht geven in het inkoopproces en omzetten. Deze knelpunt zullen deels, maar niet geheel, worden opgelost met de ingebruikname van het nieuwe rapportagesysteem in Power BI.

6. Wat is de tolerantie voor lage datakwaliteit?

Er is nauwelijks tolerantie voor lage datakwaliteit. Als dit eenmaal wordt geconstateerd zullen veel gebruikers besluiten geen gebruik te maken van de data en terug te vallen op “gut-feeling”.

7. Welke governance is er om de verbetering van de gegevenskwaliteit te ondersteunen?

Binnen Croonwolter&dros is er nog geen formele governance structuur geïmplementeerd. Dit is wel de wens van management en er zijn formele initiatieven gepland maar nog niet in uitvoering.

Deze initiatieven betreffen het opzetten van een centrale data-catalogus met o.a. definities, het vertrekken van kwaliteitslabels (“officieel”) aan rapportages en het aanwijzen van data-stewards/-eigenaren

8. Welke aanvullende governance structuren zijn nodig, om de datakwaliteit te waarborgen?

Zie 7.

5.1.6.2 Datakwaliteit strategie

De datakwaliteit strategie beschrijft waarom datakwaliteit management van belang is binnen de organisatie en wat voor strategie er gekozen is om de datakwaliteit te managen. Dit hoofdstuk is onderdeel van “Hoe wordt de datakwaliteit getoetst?”. Het eerste deel van dit hoofdstuk beschrijft het belang van datakwaliteit voor de inkoopafdeling, met een voorbeeld vanuit de praktijk. Vervolgens wordt de aanpak van de datakwaliteitstool beschreven. Tenslotte wordt er uitgelegd waarom data stewards van belang zijn voor datakwaliteit management en waar zij verantwoordelijk voor zijn.

Croonwolver&dros heeft meer dan 400 raamcontracten met leveranciers. Elke leverancier heeft zijn eigen specialiteiten en kenmerken. Naast de kenmerken heeft elke leverancier ook bepaalde bonussen, targets en kortingsafspraken in zijn raamcontract. Croonwolver&dros kan veel korting halen aan het einde van het jaar, indien de bonustargets zijn gehaald. Een voorbeeld van een bonustarget is: *Als Croonwolver&dros bij leverancier X voor 3 miljoen euro aan materiaal besteld, krijgt de organisatie 5% extra korting over de omzet.* Met deze enkele afspraak met de leverancier kan de organisatie €150.000 besparen.

Bij het maken van afspraken wordt onderhandeld tussen de organisatie en de leveranciers. Tijdens onderhandelingen met een leverancier is het van belang dat de gegevens die getoond worden kloppen met de werkelijkheid. Informatie over de leveringsperformance, betaaltermijnen, aantal openstaande betalingen, klachten of opmerkingen kunnen invloed hebben op de onderhandelingen.

Een aantal voorbeelden deze invloeden kunnen hebben op de onderhandelingen:

- De openstaande facturen van de leverancier zijn nog niet doorgevoerd in het systeem, waardoor de totale omzet van de leverancier niet de werkelijkheid weergeeft.
- De betaalde facturen worden weergegeven met het totaal aan omzet, maar er is geen onderscheid gemaakt tussen te laat betaalde facturen en op tijd betaalde facturen.
- De facturen van leverancier X wijken 80% van de tijd af van de originele bestelling. Dit kun jij zien voordat je in onderhandeling met de leverancier gaat, omdat je datakwaliteit up-to-date is en de integriteit en betrouwbaarheid hoog is.
- Je kunt de afspraken die zijn gemaakt in de vorige raamcontracten inzien, waardoor je een patroon of een trend kunt herkennen.
- De aantekeningen die bij de leveringen van de leverancier zijn genoteerd door de gebruikers (projectleiders, inkopers), waardoor je inzicht krijgt in de leveranciersperformance.

De bovenstaande voorbeelden kunnen een negatief of positief effect hebben op de onderhandelingen. Wanneer je onderhandeld wil je zoveel mogelijk weten over de leverancier, om een zo goed mogelijke deal te krijgen en niet voor verrassingen komen te staan.

De gegevens in de tabellen hebben verschillende soorten waarden. Er is transactie data, persoonlijke data, referentie data, maar niet alle velden hebben evenveel invloed op de business. Om deze verschillen te identificeren en te prioriteren is er met de stakeholders naar de data gekeken en zijn er verschillende soorten “gewichten” gedefinieerd. De gewichten zijn gerangschikt van 3: hoogste prioriteit naar 0: laagste prioriteit. De gewichten worden meegenomen in de datakwaliteit beoordeling en staat in relatie tot een kolom.

Gewicht	Impact	Prioriteit
3	De data is cruciaal bij het rapporteren, de resultaten waar de organisatie op stuurt, worden beïnvloed door deze gegevens.	Hoogste prioriteit, hier dient een melding van gemaakt te worden naar de data steward(s). Dit moet z.s.m. opgelost worden. Dit dient binnen een week opgelost te worden.
2	Wordt gebruikt in het dashboard, maar is niet cruciaal bij het rapporteren op kengetallen.	Middel prioriteit, de business wordt niet direct aangetast. Maar functionaliteiten gebruiken wel de informatie. Dien een melding in om dit binnen 2 weken op te lossen.
1	De data wordt gebruikt als context bij het rapporteren, maar heeft geen invloed op de business.	Lage prioriteit, Dit kan de leesbaarheid aantasten van de rapportage. Indien het meer baten heeft dan kosten, dien een melding in.
0	Wordt niet gemeten, omdat het geen zakelijke impact heeft.	Dit wordt niet gemeten, dus geen prioriteit.

Tabel 8 data gewichten

Er is een formule gekozen om de beoordeling in één getal weer te geven. Bij het beoordelen van datakwaliteit is het vereist dat dezelfde input, altijd dezelfde output geeft. Met deze formule wordt de uitkomst gestandaardiseerd en kan op iedere tabel toegepast worden. De formule is als volgt:

Het gewogen gemiddelde ($\bar{x}$) is gelijk aan de som van het product van het gewicht (w_i) maal het gegevensaantal (x_i) gedeeld door de som van de gewichten.

$$\bar{x} = \frac{\sum_{i=1}^n w_i \cdot x_i}{\sum_{i=1}^n w_i} = \frac{w_1 x_1 + w_2 x_2 + \dots + w_n x_n}{w_1 + w_2 + \dots + w_n}$$

Figuur 8 Datakwaliteit formule

Voorbeeld: Leverancier data scoort op de volgende velden:

Kolom	Aantal rijen	Nullcount	Gewicht
Postcode	1000	50	3
Huisadres	1000	50	3
Naam	1000	75	2
E-mailadres	1000	100	1

Tabel 9 voorbeeld formule

Wanneer dit wordt toegepast met de formule:

$$(0.05 \cdot 3 + 0.05 \cdot 3 + 0.075 \cdot 2 + 0.1 \cdot 1) / (3 + 3 + 2 + 1) = 0.55 / 9 = 0.061$$

Dit getal kan dan vervolgens omgezet worden naar percentage: $1 - 0.061 \cdot 100 = 93.9\%$

Hieruit kan geconcludeerd worden dat de leveranciers data voor 93.9% compleet is.

De volgende stap in het opzetten van de datakwaliteitscontrole is het doorzoeken van de data en de bevindingen noteren. De bevindingen worden met de data steward doorgenomen. Het belang hiervan is om de data te benaderen vanuit een ander perspectief, wat kan leiden tot nieuwe inzichten, vragen of antwoorden.

Tabel	Kolom	Gewicht	Data rules	Bevindingen
-------	-------	---------	------------	-------------

Tabel 10 datakwaliteit aspecten meta data

De bedrijfsregels en datakwaliteitsnormen worden gedefinieerd op basis van de bedrijfsvereisten. De data rules worden vastgelegd samen met de data steward en mogelijk met professionals die kennis hebben van de data en de business. In de data rules wordt bepaald wanneer een veld goed of fout is. Het gewicht wordt bepaald samen met de data steward en de professionals. In [Tabel 10 datakwaliteit aspecten meta data](#) staat het eindresultaat in de vorm van meta data van het opstellen van de data rules met de gewichten.

5.1.6.2.1 Data stewards

Data Stewards beheren de data in het belang van de organisatie en de belanghebbende, zij moeten ervoor zorgen dat bedrijfsgegevens van hoge kwaliteit zijn en effectief kunnen worden gebruikt. Data Stewards zijn ook verantwoordelijk voor data governance activiteiten. Voor elke afdeling dient er een Business Data Steward benoemd te worden om de datakwaliteit te waarborgen en te beheren. Een data steward bewaakt de datakwaliteit en wordt vanuit een Data Governance project aangesteld.

Voor de inkoop data marts dient een Business Data Steward benoemd te worden. Een Business Data Steward is een zakelijke professional, in dit geval is een inkoper verantwoordelijk voor de datakwaliteit van de inkoopafdeling. De inkoper werkt samen met belanghebbende en gebruikers om de data te beheren en om definities vast te leggen. De data steward past niet de foutieve data aan, maar stuurt een rapport naar de data beheerders van zijn/haar afdeling die de wijzigingen doorvoeren.

5.1.6.3 Eerste datakwaliteit beoordeling

Voor de eerste datakwaliteitscontrole is er handmatig een controle gedaan in Excel met behulp van Power BI. Deze meting is gebaseerd op de data rules die gedefinieerd zijn. De reden om dit handmatig te doen is om het later, wanneer dit wordt toegepast met scripts en andere tools of functionaliteiten dat het vergeleken met elkaar kan worden. Hier is een risico aan gebonden. Indien de resultaten uit Power BI en Excel niet overeenkomen met de resultaten van de geïmplementeerde tool, moeten de afwijkingen geïdentificeerd en verbeterd worden.

De eerste beoordeling is uitgevoerd op de tabellen van leverancier en leverancierskenmerken. De leverancier gerelateerde data is één van de meest gebruikte data onderdelen door inkopers. Inkopers hebben dagelijks te maken met leveranciers, voor bestellingen, afspraken of onderhandelingen. Het doel van deze beoordeling is om een betrouwbaar proof of concept op te leveren die de huidige datakwaliteit status van de leverancier tabellen weergeeft.

Het datakwaliteit management tool is in SQL Server ontwikkeld. Dit wordt gedaan doormiddel van scripts (Zie [Bijlage 4 Datakwaliteit](#)) die over alle tabellen en kolommen itereert die van belang zijn voor de business. Dit wordt bepaald samen met de data steward en business professionals. Data met gewicht "0" worden niet meegenomen in de datakwaliteit meting.

Het script in SQL Server is op een dusdanige manier opgezet, dat het snel en efficiënt uitgebreid kan worden naar andere tabellen. De resultaten van de Excel-meting en de SQL Server zijn in Figuur 9 weergegeven. In de tabel zijn de overeenkomsten te zien tussen de twee werkwijzen. Meer informatie over het meetinstrument in Excel is terug te vinden in [10.5 Bijlage 5 meetresultaten Excel en Power BI](#). De resultaten in SQL Server zijn weergegeven in [10.6 Bijlage 6 meetresultaten SQL Server](#).

tabel	kolom	Aantal nullwaarden EXCEL	Aantal foutieve waarden EXCEL	Aantal dubbele waarden EXCEL	SQL Server aantal foutieve waarden
d_leverancier	Leverancier_skey	0	0	0	0
d_leverancier	Administratie	0	0	0	0
d_leverancier	Bestelmedium	0	0	6309	6309
d_leverancier	Betalingsconditie	0	0	704	704
d_leverancier	Betalingsconditie_crediteur	2848	0	17	2865
d_leverancier	Betalingsstermijn_crediteur_dagen	0	10007	0	10007
d_leverancier	Contactpersoon	-	-	-	-
d_leverancier	Ean_adres	-	-	-	-
d_leverancier	Emailadres	87	0	0	87
d_leverancier	Inkoopconditie	1568	157	2030	-
d_leverancier	Kernleverancier	37	0	0	37
d_leverancier	Leveranciers nr	-	-	-	-
d_leverancier	Leveringsconditie	0	5841	0	5841
d_leverancier	Naam	2835	5	0	2840
d_leverancier	Orderkosten	-	-	-	-
d_leverancier	Status	0	0	0	0
d_leverancier	Telefoonnummer	-	-	-	-
d_leverancier	Wka_leverancier	0	0	0	0
d_leverancierkenmerken	Leverancierkenmerk_skey	-	-	-	-
d_leverancierkenmerken	Admin_code	0	0	0	0
d_leverancierkenmerken	Leveranciercode	-	-	-	-
d_leverancierkenmerken	Leverancier_skey	-	-	-	-
d_leverancierkenmerken	Labelcode	0	0	0	0
d_leverancierkenmerken	Labelvolgnummer	-	-	-	-
d_leverancierkenmerken	Soort	0	0	0	0
d_leverancierkenmerken	Kenmerk	1758	0	0	1758
d_leverancierkenmerken	Omschrijving	12	0	0	12
d_leverancierkenmerken	Waarde	10576	0	0	10576
d_leverancierkenmerken	Datum geldig van	4129	0	0	4129
d_leverancierkenmerken	Datum geldig tot en met	386	0	0	386

Figuur 9 vergelijking Excel en SQL Server meting

5.1.6.4 *Verbeteringen identificeren en prioriteren*

De prioriteiten van de velden zijn bepaald aan de hand van de gewichten die aan de velden zijn gehangen. Met deze manier van prioriteren kunnen de gebruikers efficiënt inzicht krijgen op de data velden die een impact hebben op de business.

Het kan voorkomen dat er foutieve data in de tabellen staan waar Croonwolder&dros geen invloed op heeft. Acto is de leverancier van de leverancier-data en levert dit aan Croonwolder&dros. Er zijn voorvallen waar data verkeerd in Acto staat opgeslagen. Indien dit voorkomt, gebeurt er het volgende:

- Probleem identificeren
- Probleem prioriteren
- Hoeveelheid gegevens beïnvloed noteren
- Aantal en type bedrijfsprocessen dat door het probleem wordt beïnvloed
- Risico's verbonden aan het probleem
- Kosten van het herstellen van het probleem en/of de oorzaken

In het hoofdstuk “Welke opmerkelijke datakwaliteit gerelateerde bevindingen zijn er gevonden?” Staan de bevindingen beschreven.

5.1.6.5 *Definieer doelen voor verbetering van datakwaliteit*

De doelen kunnen worden gedefinieerd op operationeel, tactisch en strategisch niveau. De leverancier-doelen zijn gedefinieerd op tactisch niveau. Het doel is om **98%** van de gebruikte leveranciers volledig correct in het datawarehouse te hebben staan. Dit doel is door de data steward gedefinieerd. Voor iedere data mart dienen de data stewards samen met de stakeholders en het management te definiëren wat de doelstellingen zijn omtrent de datakwaliteit.

5.1.6.6 Datakwaliteit modificaties ontwikkelen en implementeren

De resultaten van de datakwaliteit meting zijn in een dataerror en datacontrole tabellen geplaatst in SQL Server. Deze tabel kan via de SQL Server ingeladen worden in Power BI. In Power BI zijn relaties gemaakt tussen de originele leverancier en leverancierkenmerken tabellen.

_Dataerror

De tabel dataerror is om de waarden met de foutieve data te kunnen herleiden naar de originele tabellen. Een voorbeeld: *Leverancier X met leveranciersnummer 101 besteld via de mail. In de dataset van Croonwolter&dros heeft leverancier X een postcode staan in plaats van een e-mailadres. In de dataerror tabel staat bij "Omschrijving" 101 en in de kolom "Waarde" staat de waarde "1011AX".*

Met het leveranciersnummer kan herleid worden om welke leverancier het gaat.

De dataerror tabel bevat de volgende meta-data:

Kolommen	Omschrijving
Tabel	De tabel waar de data error in gevonden is
Kolom	De kolom waar de data error in gevonden is
Code	Administratie voor inkoop, kan een andere referentiele waarde zijn voor andere afdelingen. Op deze waarde kan gefilterd worden. Hier moet een waarde in komen die van belang is voor de business.
Conditie	Welke condities zijn gebruikt voor het bepalen of de data fout is ja/nee
Omschrijving	In de omschrijving staat de surrogate key voor de inkoopafdeling. De waarde die in dit veld staat is de primaire sleutel in de tabel waar de data error voorkomt.
Waarde	De waarde die het veld inhoudelijk bevat.

Tabel 11 _dataerror

_Datacontrole

De Datacontrole tabel bevat de volgende meta-data:

Kolommen	Omschrijving
Tabel	De tabel waar de data error in gevonden is
Kolom	De kolom waar de data error in gevonden is
Conditie	Welke condities zijn gebruikt voor het bepalen of de data fout is ja/nee
Aantal	Het aantal data errors de kolom bevat

De Datacontrole tabel geeft de statistische resultaten van de tabellen. Deze tabel is niet te combineren met de originele tabellen van bijvoorbeeld leverancier of leverancierkenmerken. Waarom is dat zo? De tabel weergeeft de aantallen weer en kijkt dus niet per regel naar de data, maar per statement hoeveel regels overeenkomen aan de voorwaarden.

Hoe weet je nu zeker dat de resultaten in beide tabellen hetzelfde zijn? De tabellen gebruiken dezelfde statements in het script, deze komen dus in alle gevallen overeen met elkaar.



Figuur 10 relaties datakwaliteit tabellen

Tabel	Functie
D_raamcontract	Verborgen relatie, heeft geen functie binnen dit project.
_data_error_leverancier	Bevat alle foutieve data van de tabel leverancier. De kolom Waarde bevat tekst/getallen/nulwaarde wat fout is volgens de data rules.
D_leverancier	De tabel met alle informatie over de leveranciers en op welke manier de leverancier besteld.
Leverancier_raamcontract	Verborgen relatie, heeft geen functie binnen dit project.
_datacontrole	Verouderde versie van de datacontrole, niet relevant meer. Deze tabel is vervangen door _datacontrole_administratie.
_dataerror	Bevat alle velden met foutieve waarden, deze is opgesplitst in _data_error_leverancier en _data_error_leverancierkenmerken.
_datacontrole_administratie	Met deze tabel wordt de formule uitgevoerd om de eindscore te berekenen. Deze staat in relatie met _dataerror om het rapport interactief te maken.
D_leverancier_gewicht	In deze tabel zijn de gewichten van de velden gedefinieerd.
D_leverancierkenmerken	De tabel met alle informatie over de kenmerken van elke leverancier. Contracten, performance, administraties, enzovoort.
_data_error_leverancierkenmerken	Bevat alle foutieve data van de tabel leverancierkenmerken. De kolom Waarde bevat tekst/getallen/nulwaarde wat fout is volgens de data rules.

Figuur 11 uitleg inhoud tabellen leverancier data



Het SQL script waarmee de resultaten in een tabel worden geschreven is volledig geautomatiseerd en kan op elk moment aangeroepen worden. Het script kan op elk gewenst tijdstip uitgevoerd worden.

In Power BI is de formule (Zie Figuur 8 Datakwaliteit formule) toegepast op de tabellen _dataerror & _datacontrole_administratie. De DAX Statements die zijn gebruikt voor het berekenen van de datakwaliteit gestandaardiseerde score, zijn terug te vinden in de bijlage ([Bijlage 7 DAX Statements formule gestandaardiseerd datakwaliteit cijfer](#)).

De formule is gevalideerd met de BI Manager. Op verschillende manieren is dezelfde formule gebruikt en iedere tool of methode komt uit op dezelfde eindscore. Hiermee kunnen we concluderen dat de formule klopt. De test die is weergegeven is gebaseerd op de leverancierkenmerken datakwaliteit. De visualisatie van de datakwaliteit score is terug te zien in het hoofdstuk “

Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?”.

Weighted Average Calculator

Weighted average calculator and calculation. Weighted mean calculator.

Weight	Data number
2	0.0053
2	0.0572
1	0.0243
1	0.0002
1	0.1464

Weighted average:
0.04227143

Total of weights:
7

Calculation:
$$(2 \times 0.0053 + 2 \times 0.0572 + 1 \times 0.0243 + 1 \times 0.0002 + 1 \times 0.1464) / (2 + 2 + 1 + 1 + 1) = 0.04227143$$

Figuur 12 validatie formule datakwaliteit

De uitkomst van de online gewogen gemiddelde rekenmachine geeft een “Weighted average” = 0.04227. Dit is het foutpercentage. De eis is om het correcte percentage te weergeven, dus wordt het: $1 - 0.04227 = 0.95773$. Als dit naar de eenheid percentage wordt berekend wordt de uiteindelijke eindscore: $0.95773 * 100\% = 95.77\%$.

5.1.7 Welke tools kunnen gebruikt worden voor datakwaliteit management?

Voor deze deelvraag zijn de mogelijkheden onderzocht om de datakwaliteit tool in te ontwikkelen. Dit wordt gedaan op twee verschillende manieren:

- Er is onderzoek gedaan naar de verschillen tussen zelf een datakwaliteit tool ontwikkelen vs. aanschaffen.
- Verschillende tools zijn onderzocht om de juiste tool voor de juiste functionaliteiten te kiezen.

5.1.7.1 Datakwaliteit tool zelf ontwikkelen vs. aanschaffen

Croonwolver&dros wil alle mogelijkheden onderzoeken om de best passende oplossing voor de organisatie te implementeren. De oplossing moet duurzaam, betaalbaar en beheersbaar zijn, naast dat het de datakwaliteit kan meten op basis van de business. Eind 2020 wordt er een budget opgesteld die wordt gereserveerd voor Datakwaliteit management. Deze deelvraag geeft antwoord op de voor- en nadelen van het ontwikkelen van een eigen tool, of het aanschaffen van een tool van een externe organisatie.

Dataladder (DataLadder, 2019) is de tool die vergeleken wordt met het ontwikkelen van een eigen datakwaliteit tool. DataLadder heeft een product "DataMatch Enterprise" die over de requirements van de organisatie beschikt (Zie Tabel 17).

Prioriteit	Requirement
Must have	Data matching, één veld dat anders geregistreerd is door afwijkende karakters, clusteren tot één.
Must have	Data regels integratie, het zelf definiëren van data regels.
Must have	Onder de €10.000,- per jaar
Must have	Rapportages maken
Must have	Centrale database voor definities (Data catalog)
Must have	Log bijhouden zodat er analysis op historische datakwaliteit vs. huidige datakwaliteit kunnen worden uitgevoerd.
Should have	Exporteren naar Visualisatietools (Power BI)
Would have	RBAC (Role Based Access Control wordt in Azure beheerd)
Would have	Real-Time data

Tabel 12 datakwaliteit tool requirements

DataLadder versus eigen datakwaliteit tool

In een online demo heeft een data analist de mogelijkheden van het product DataMatch Enterprise laten zien. Uit deze demo is een vergelijkingsonderzoek toegepast.

DataLadder DataMatch EnterPrice	
Voordelen	Nadelen
Kan snel geïmplementeerd worden	Hoge kosten
Veel functionaliteiten	Niet mogelijk om uit te bouwen naar eigen business
Betrouwbare functionaliteiten	Enkel op één desktop te gebruiken, anders lopen de kosten op
Geavanceerde functionaliteiten	Ondersteuning enkel via telefoon, email of skype
Business users kunnen het gebruiken	Extra kosten gebonden aan aanpassingen implementeren naar eigen behoeften
Kan exporteren naar Excel of visualisatietools	Kans op veel ongebruikte functionaliteiten waar wel voor wordt betaald
Functionaliteiten op maat laten maken	Elke aanpassing/extra functie kost geld
Geen onderhoud aan de tool	

Tabel 13 vergelijkingsonderzoek Dataladder

Croonwolver&dros doet op het moment nog niet gestructureerd aan datakwaliteit management. Een product zoals de DataMatch Enterprise bezit alle must haves van de requirements, maar het is een er zitten wel hoge kosten aan verbonden. De tool kost €10.000 per jaar. Voor dit geld kan het product maar op één desktop gebruikt worden. Voor veel functionaliteiten betaal je, maar worden in de praktijk niet of nauwelijks gebruikt.

5.1.7.2 Datakwaliteit tools

5.1.7.2.1 Power BI versus SQL Server

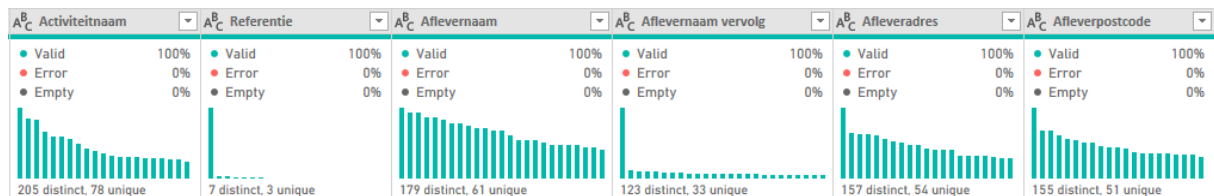
Om te kunnen kiezen tussen de twee tools welke het beste geschikt is, dient het doel van het product duidelijk te zijn en welke tool voor welke doeleinden gebruikt wordt. Het doel van de inkoopafdeling voor deze tool is: *“De inkoopafdeling wil een datakwaliteit tool die aansluit op de bedrijfsprocessen, waardoor eindgebruikers real-time de datakwaliteit kunnen valideren en wijzigingen doorvoeren. Dit met doel om de bedrijfsprocessen te ondersteunen in het maken van beslissingen.”*

Power BI

De activiteiten die leiden tot het product kan op verschillende manieren ontwikkeld worden. Microsoft heeft sinds Oktober 2018 een [update](#) doorgevoerd binnen Power BI die het mogelijk maakt om in enkele minuten data profiling toe te passen op kolommen in een tabel. Deze tool weergeeft het aantal correcte (valid) gegevens, het aantal errors en blanks. Op het eerste opzicht lijkt dit een enorm handige tool waarmee je snel de status van je kolommen kan inzien. Echter zitten er minpunten aan deze tool:

- De Data Profiling tool is niet bewerkbaar. Hierdoor kun je niet de tool aanpassen naar de wensen en eisen van de organisatie. Vervolgens is het niet mogelijk om verder te bouwen met de tool, waarmee het erg gelimiteerd blijft.

- De *Empty* waarden zijn niet (altijd) correct, de *empty* waarde is gedefinieerd op basis van de regels van Power BI, de definitie van “leeg” kan verschillen tussen organisaties en tabellen.
- Het weergeeft enkel de status op drie verschillende aspecten; Valid, Error en Empty. Met het valideren van de gegevens is het van belang dat er verder gekeken wordt dan enkel naar deze drie aspecten.
- Het kost veel tijd. Power BI's verwerkingssnelheid verminderd drastisch wanneer het met tekstvelden werkt. Indien dit over miljoenen rijen gedaan wordt, kost elke activiteit zoals een filter of een tabel toevoegen veel tijd.



Figuur 13 Power BI Data Profiling

De eerste keuze om de datakwaliteit tool in te ontwikkelen was Power BI. Tijdens het implementeren kwamen er steeds meer knelpunten bij kijken. De voor- en nadelen van Power BI zijn hiervoor bijgehouden:

Voordelen Power BI

- Power BI kan itereren over elke tabel en kolom om de Data Profiling en Data Schema waarden op te halen. Dit kan met een enkel script.
- De tabellen zijn makkelijk te transformeren naar de wensen van de gebruiker (mits de gebruiker kennis van Power BI heeft).
- Stappen worden bijgehouden en kunnen met gemak weer verwijderd worden.
- Fuzzy Matching is een functie binnen Power BI.
- Het is een low code tool. Mensen zonder professionele IT-vaardigheden kunnen analyses uitvoeren op de data.

Nadelen Power BI

- De data rules van Croonwolter&dros zijn niet toegankelijk in de Data Profiling functies, dit dient in DAX te gebeuren en dat kost veel handmatig werk en daarmee wordt het onbeheersbaar. Het aantal measures dat gemaakt moet worden wordt moeilijk te beheren en te onderhouden.
- Aanpassingen maken op grote hoeveelheden aan data kost veel tijd om binnen Power BI uit te voeren. Dit gaat ten koste van de performance.
- Hoe meer functionaliteiten in Power BI wordt toegevoegd, hoe langer elke wijziging duurt. De tool vereist veel RAM-geheugen. Een measure creëren kan al tot 10 seconden kosten. Hierdoor wordt het onmogelijk om met Power BI te werken.
- Data Matching (Zie Figuur 17 Data Matching 1) is lastig te beheren op grote datasets, het is een handmatig proces. Met grote datasets is dit onmogelijk te managen.
- Iedere eindgebruiker kan

SQL Server

Met het implementeren van de activiteiten in SQL Server is er meer vrijheid. Het kost tijd om de juiste manier te vinden om de tool op te stellen. Tijdens het ontwikkelen van de datakwaliteit tool

kwamen er een aantal knelpunten bij kijken. De voor- en nadelen van SQL Server zijn hiervoor bijgehouden:

Voordelen SQL Server

- Vrijheid in het ontwikkelen
- Performance is hoog, ook bij grote hoeveelheden data
- De eindgebruiker kan niet bij de originele scripts

Nadelen SQL Server

- De tabellen hebben een vaste structuur en kunnen niet door de eindgebruiker aangepast worden
- Wijzigingen in de tool kunnen enkel door een IT-professional doorgevoerd worden. Dit kan veel tijd kosten.
- Heeft geen tools, alles moet zelf ontwikkeld worden. Dit kost veel tijd
- Eindgebruiker kan geen validatie doen en krijgt alleen de resultaten te zien

Met allebei de tools is het mogelijk om een datakwaliteit tool te ontwikkelen. Maar wat is de meest toegankelijke manier om die te doen? Hiervoor wordt er gekeken naar het doel waarvoor je welke tool wil gebruiken. Power BI wordt voornamelijk gebruikt door Business Professionals die over het algemeen minder IT-vaardigheden hebben dan een database administrator. Deze professionals willen rapportages kunnen creëren met de beschikbare data. SQL Server is een tool die gebruikt wordt door IT-professionals. SQL Server wordt gebruikt voor het configureren, beheren en analyseren van de data. Het is kan doormiddel van query's relaties creëren tussen tabellen in de database en vervolgens hier analyses op loslaten. Het kan de resultaten van de analyses in een tabel schrijven.

Uit de resultaten komt voort dat SQL Server de beste oplossing is om de activiteiten in te implementeren. Het configureren van de data en de functionaliteiten worden in SQL met query's gebouwd. De resultaten worden in tabellen geschreven. Vervolgens visualiseert Power BI de resultaten die voortvloeien uit de SQL query's. Hiermee voorkom je dat Power BI traag wordt, dit verhoogd de gebruikersvriendelijkheid van de tool. De tabellen zijn een vaste structuur, dus wijzigingen in de data rules komen niet vaak voor.

5.1.7.2.2 Microsoft SQL Server Data Quality Services

Met de migratie naar een Cloud omgeving is onderzoek uitgevoerd naar een service binnen Microsoft Azure die de datakwaliteit kan meten en managen. Binnen Microsoft Azure is er geen service voor datakwaliteit management. Dit dient een organisatie zelf op te zetten in de Data Factory of met een tool buiten Azure te integreren aan het ETL proces.

Microsoft heeft een service SQL Server Data Quality Services (DQS). DQS bevat alle domeinen die de organisatie stelt voor een datakwaliteit management tool:

- **Data Cleansing:** Wijzigen, verwijderen of verrijken van gegevens die onjuist of onvolledig zijn.
- **Data Matching:** Wordt gebruikt om data sets met elkaar vergelijken, koppelen en om dubbele waarden te verwijderen.
- **Profiling:** Profileren van de data sets. Biedt inzicht in de kwaliteit van de gegevens, bijvoorbeeld minimale en maximale waarden, het gemiddelde of de standaarddeviatie.
- **Monitoring:** Het volgen van en bepalen van de status van activiteiten voor datakwaliteit. Met monitoring kun je bijhouden of de tool doet waar het voor ontworpen is.
- **Knowledge base:** Eigen business- en data rules definiëren en integreren in de tool. De tool analyseert de data op basis van de gedefinieerde regels, waarmee de datakwaliteit efficiënt en doelgericht verhoogd wordt.

De tool kan niet gebruikt worden in de nieuwe Cloud omgeving. Microsoft Azure heeft deze tool (nog) niet in Azure beschikbaar gemaakt. DQS is enkel beschikbaar voor SQL Servers op Windows computersystemen. Dit is een nadeel van Microsoft Azure, je bent afhankelijk van de services die Microsoft Azure beschikbaar stellen. Om datakwaliteit management te integreren voor Croonwolter&dros moet er een externe datakwaliteit management tool aangeschaft worden of de organisatie ontwikkelt zelf een tool.

5.1.8 Welke opmerkelijke datakwaliteit gerelateerde bevindingen zijn er gevonden?

Acto (Acto, 2019) is het ERP systeem van Croonwolver&dros. Dit systeem verwerkt de logistieke, administratieve en financiële bedrijfsprocessen van Croonwolver&dros in één informatie- en managementsysteem.

Met het ontwikkelen van een nieuw dashboard voor de inkoopafdeling en het implementeren van een datakwaliteit tool, is er een probleem geïdentificeerd met de projectnummers. Figuur 14 omzet per project is een snapshot van het dashboard dat in productie genomen wordt. In dit figuur is te zien dat meer dan 630 miljoen euro niet aan een project staat gerelateerd. Hiermee is het analyseren per project niet betrouwbaar, ook wel onmogelijk. De oorzaak hiervan is niet bekend en de oplossing dient opgelost te worden door Acto.

Contract overzicht per Project		
Project naam	DER CONTRACT	Totaal
11962001 - DSM REGIENUMMER 1E KWARTAAL 2016	€ 572.426.469	€ 632.242.497
11962003 - DSM REGIENUMMER 3E KWARTAAL 2016	€ 281	€ 281
11962004 - DSM REGIENUMMER 4E KWARTAAL 2016	€ 801	€ 801
11962005 - DSM, MJOP 2016 OFF-000-901-0-0.E.005	€ 9.421	€ 9.421
11962012 - DSM, GARANTIE MJOP 2015	€ 328	€ 328
11962032 - LMA 2016 DSM GEB EXPLOITATIE G122-051	€ 1.593	€ 1.593
11962033 - LMA 2016 DSM GEB EXPLOITATIE G122-093	€ 77	€ 77
11962034 - LMA 2016 DSM GEB EXPLOITATIE G122-059	€ 726	€ 726
11962036 - LMA 2016 DSM GEB EXPLOITATIE G122-094	€ 1.247	€ 1.247
11962039 - LMA 2016 BRIGHTLANDS GEB G124-037	€ 2.651	€ 2.651
11962040 - LMA 2016 DSM GEB INVESTERING G122-094	€ 315	€ 315
11962041 - DSM, MJOP 2016 OFF-000-901-0-0.E.005	€ 670	€ 670
Totaal	€ 689.069.412	€ 752.259.737

Figuur 14 omzet per project

Onderstaande figuur is een bestelorder in Acto geselecteerd met meerdere (6) bestelorderregels. De geselecteerde bestelorderregel heeft geen projectnummer in tegenstelling tot alle andere bestelorderregels. Dit probleem komt in meerdere gevallen voor.

Zoeken in 'Wijzigen en verzenden bestelorders (LPM201 / PBATEN / PBATEN_201 / abprod1 / P / B110 / 19.10)', alternatief=INLEEN+										
Batchnr.	LVA	Leverancier	Bestelorde	Bestelp	Bestelreg	Projectnumme	Vestigti	vestiging	Aanvrager	Bruto bedrag
81912001	PMC ZYP	23294443	01	05	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.	66,96	27/03/2019 3/Geprint
81912001	PMC ZYP	23294443	01	04	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.	144,77	27/03/2019 3/Geprint
81912001	PMC ZYP	23294443	01	06	00000000	400	IND-Divisie Industrie	Zeeuw de M.S.	247,26	27/03/2019 3/Geprint
81912001	PMC ZYP	23294443	01	02	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.	1.287,80	27/03/2019 3/Geprint
81912001	PMC ZYP	23294443	01	01	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.	2.833,16	27/03/2019 3/Geprint
81912001	PMC ZYP	23294443	01	03	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.	128,78	27/03/2019 3/Geprint

Figuur 15 bestelorder scherm acto

De projectnummers moeten ergens geregistreerd staan in Acto. Dit is een verplicht veld van het ERP systeem (Zie Figuur 16 projectnummers en kostenplaats registratie ACTO). Dit is terug te zien in een ander scherm van Acto. In Figuur 16 staat een projectnummer en een kostenplaats ingevoerd. Hieruit kan dus geconcludeerd worden dat ergens in het proces van wegschrijven van de data een consistentiefout zit. Van dit probleem zijn de snapshots naar Acto verstuurd om dit op te lossen in het systeem.

Inkoopfacturen (INK150 / PBATen / PBATen_201 / aikprod / P / B110 / 19.10)																																																																		
Bestand Start Bewerken Zoeken Extra Speciaal Vensters Help																																																																		
PgUp		PgDn		Gaanar		Toevoegen		Uitbreiden		Wijzigen		Kolom		Verwijder		Rapporten		Archiveren		Samenvoegen		Zoeken		Vorige		Eerstvolgende		Zactualiseren		Zoekbox...		Zvorige		Zeerstvolgende		Taakmonitor																														
Creditur				464443 Interparking Nederland B.V.												Status				Verwerkt																																														
Boekstuknummer				01010744												Bestelorder(s)				00000000 Besteld																																														
Factuurnr lev				0002200299												Factuurdatum				01/01/2019 Ontvangen																																														
Factuurbedrag				503,41 EUR												Geboekt bedrag				416,04 Gefactureerd																																														
BTW-bedrag				87,37 EUR												Noo te boeken				0,00 Oopenstaand																																														
1 Algemeen				Opdracht																																																														
2 Regels				<table><tr><th>Marker</th><th>Rg nr</th><th>Bestel order</th><th>Best post</th><th>Beste regel</th><th>Boeking soort</th><th>Grootboek nummer</th><th>Artikel nummer</th><th>Omschrijving</th><th>Aantal factuur</th><th>Bedrag factuur</th><th>Aantal besteld</th><th>Aantal ontvangen</th><th>Aantal gefactureerd</th><th>Project nummer</th><th>Pe nu</th><th>Co inv</th><th>Activiteit/ Proj. deel</th><th>Kosten soort</th><th>Kosten plaats</th><th>Budget bedrag</th><th>Bt co</th></tr><tr><td>☐</td><td>1</td><td></td><td></td><td></td><td></td><td>9302</td><td>32100</td><td>PR K Z 0/INTERPARKI</td><td>0,00</td><td>416,04</td><td></td><td></td><td></td><td>257.10016</td><td></td><td></td><td>00000</td><td>5300</td><td>00002570</td><td></td><td>03</td></tr></table>																			Marker	Rg nr	Bestel order	Best post	Beste regel	Boeking soort	Grootboek nummer	Artikel nummer	Omschrijving	Aantal factuur	Bedrag factuur	Aantal besteld	Aantal ontvangen	Aantal gefactureerd	Project nummer	Pe nu	Co inv	Activiteit/ Proj. deel	Kosten soort	Kosten plaats	Budget bedrag	Bt co	☐	1					9302	32100	PR K Z 0/INTERPARKI	0,00	416,04				257.10016			00000	5300	00002570		03
Marker	Rg nr	Bestel order	Best post	Beste regel	Boeking soort	Grootboek nummer	Artikel nummer	Omschrijving	Aantal factuur	Bedrag factuur	Aantal besteld	Aantal ontvangen	Aantal gefactureerd	Project nummer	Pe nu	Co inv	Activiteit/ Proj. deel	Kosten soort	Kosten plaats	Budget bedrag	Bt co																																													
☐	1					9302	32100	PR K Z 0/INTERPARKI	0,00	416,04				257.10016			00000	5300	00002570		03																																													

Figuur 16 projectnummers en kostenplaats registratie ACTO

Meer informatie over dit probleem is in een document vastgelegd (Zie [10.11 Bijlage 11 meeting factuurregels zonder kostenplaats](#)).

Data Matching

Data Matching is het opsporen van tekstwaarden die dezelfde betekenis hebben in de dataset, maar niet altijd identiek overeenkomen met elkaar. De datasets binnen Croonwolter&dros bevat meerdere tabellen met dit probleem. Dit gaat mis doordat verschillende administraties de velden op een verschillende manier vastleggen. In dit geval geldt ook weer dat de administratie B110 leidend is. In Figuur 17 en Figuur 18 is weergegeven welke administraties de Technische Unie anders heeft gedefinieerd dan de B110 administratie. De Technische Unie is een groothandelaar en één van de grootste leveranciers voor Croonwolter&dros. In het geval dat een inkoper of andere gebruiker op leveranciersnaam in Power BI de data filtert, zal de weergave van het totaal gefactureerde bedrag afwijken van de werkelijkheid.

In Tabel 14 is weergegeven dat het in de meeste gevallen goed gaat, maar in sommige gevallen wordt er via een verkeerde leveranciersnaam besteld. Dit soort fouten dienen worden opgelost om de data consistentie en integriteit te waarborgen en te verhogen, anders wijken de werkelijkheid en de weergave van het dashboard van elkaar af.

naam	administratie
Technische Unie	B110
Technische Unie B.V.	B923
Technische Unie	B926
Technische Unie BV	B929

Figuur 17 Data Matching 1

naam	Count of administratie
Technische Unie	17
Technische Unie B.V.	1
Technische Unie BV	1
Total	19

Figure 18 Data Matching 2

Bedrag gefactureerd	Naam
€30.673.850	Technische Unie
€169.421	Technische Unie BV

Tabel 14 Technische Unie verschil

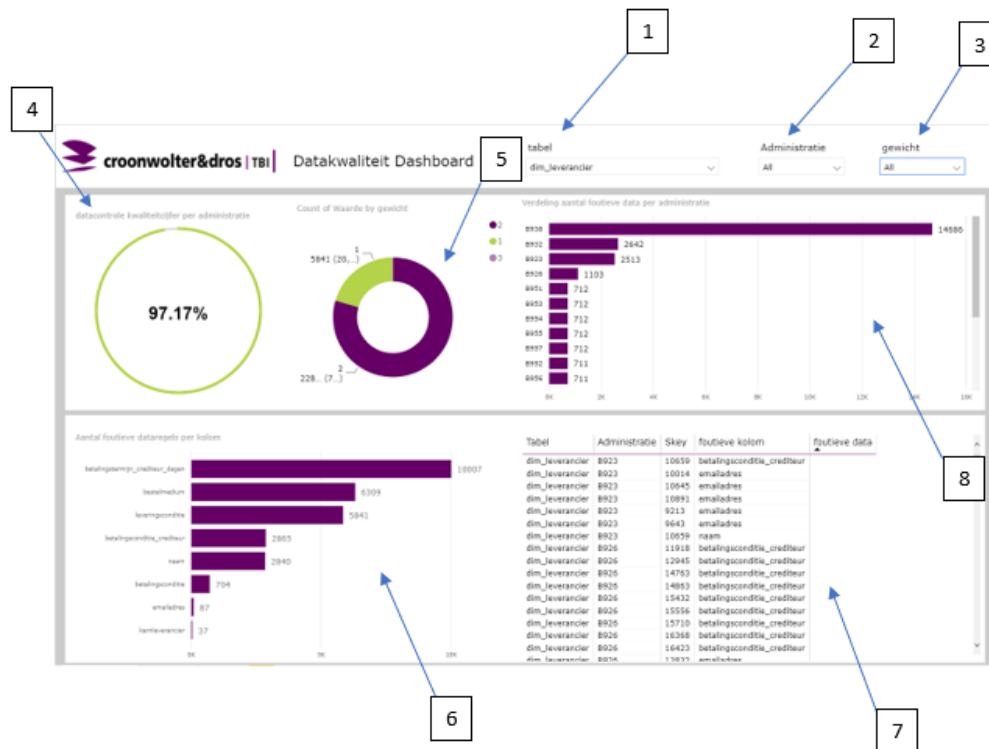
5.1.9 Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?

De databeheerders van de inkoopafdeling beheren de data. Deze business professionals hebben geen geavanceerde IT-vaardigheden. Om deze reden is het van belang dat de visualisatie van de datakwaliteit tool af te lezen is, maar ook te beheren is door mensen zonder IT achtergrond. De requirements die beginnen met “de gebruiker” zijn user requirements. De requirements die beginnen met “de tool” zijn systeem requirements. De volgende eisen zijn gesteld voor het visualiseren van de datakwaliteit:

1. De gebruiker zonder IT ervaring kan het doel van het dashboard begrijpen, zodat zij hier informatie uit kunnen halen.
2. De gebruiker kan de benodigde informatie van het dashboard aflezen, zodat hij/zij hier een vervolgactie mee kan uitvoeren.
3. De gebruiker kan het dashboard interactief bedienen, zodat hij/zij alleen de benodigde gegevens kan selecteren.
4. De gebruiker kan de tool zelfstandig beheren, zodat hij/zij de datakwaliteit kan verhogen zonder een IT-specialist in te schakelen.
5. De gebruiker dient in staat te zijn om de gegevens te exporteren naar Excel, zodat deze gedeeld kunnen worden voor bepaalde doeleinden.
6. De tool dient schaalbaar te zijn, zodat het integreren van meer data niet ten koste gaat van de datakwaliteit tool.
7. De tool dient een pagina definities te hebben, zodat onduidelijkheden voorkomen wordt.
8. De tool kan per administratie de datakwaliteit weergeven, zodat de gebruiker op administratie kan filteren.
9. De tool kan filteren op gewicht, zodat de gebruiker snel kan prioriteren welke data velden gewijzigd dienen te worden.

5.1.9.1 Visualisatie

De wensen en eisen omtrent de “look and feel” van het datakwaliteit dashboard wordt met de data steward bepaald. Het is geen moeite om in een later stadium het dashboard aan te passen naar de wensen van de gebruikers. Doormiddel van een schets op een A4 formaat is een ontwerp gemaakt (Zie [Bijlage 8 Blauwdruk schets Datakwaliteit dashboard](#)). Deze schets is vervolgens uitgewerkt in Power BI. Het ontworpen dashboard in Power BI is gevalideerd door de data steward (vertegenwoordiger van de gebruikers) en de BI manager. In (Figuur 19 DQ leverancier dashboard) is het dashboard weergegeven dat in productie is genomen.



Figuur 19 DQ leverancier dashboard

#	Soort visualisatie	Doel/Functie
1	Filter	Deze filter kan de gewenste tabel(len) filteren waarmee de datakwaliteit per tabel wordt weergegeven in het dashboard.
2	Filter	Deze filter kan de gewenste administraties filteren die de gebruiker wil weergeven in het dashboard.
3	Filter	De filter kan de gewenste gewichten weergeven van alle velden.
4	KPI Donut chart	Deze visualisatie geeft de gestandaardiseerde eindscore van de datakwaliteit van de gefilterde tabel(len) weer. Deze eindscore is berekend door de formule (Figuur 8 Datakwaliteit formule).
5	Donut chart	De donut chart weergeeft het gedeelte van het geheel weer per gewicht.
6	Horizontale staafdiagram	Deze chart geeft het aantal foutieve dataregels per kolom weer.
7	Tabel	De tabel bevat gegevens over de tabellen, kolommen, administraties, surrogate keys en de waarde die fout is. Deze tabel kan geëxporteerd worden naar Excel.
8	Horizontale staafdiagram	Het staafdiagram weergeeft de verdeling van het aantal foutieve data weer per administratie.

Tabel 15 uitleg dashboard

5.1.9.2 Definities

Definities zijn een vereist onderdeel voor een dashboard. Zonder definities is er een risico dat de informatie dat het dashboard weergeeft niet wordt begrepen door de gebruiker, maar waar worden definities opgeslagen en beheert? Als organisatie kies je er voor waar en hoe je definities bewaard en beheert. Om te waarborgen dat niet alle definities over de gehele organisatie bereikbaar zijn, worden definities binnen Croonwolver&dros per divisie, of per afdeling opgesteld en bewaard. Dit zorgt ervoor dat het beheersbaar blijft en alleen voor de juiste doeleinden wordt gebruikt. DAMA DMBOK2 (Hoofdstuk Data Model Management) schrijft het volgende over het beheren van definities:

Terms and definitions used within a source system may make sense within the confines of that system, but they do not always make sense at an enterprise level.

De definities zijn als een pagina/rapportage toegevoegd aan het dashboard in Power BI. Hiermee wordt voorkomen dat andere afdelingen of divisies deze definities kan bereiken. Het is een best practice om de definitie-pagina als laatste pagina van het dashboard te hebben staan, of als een tabel in de database.

In de definitie pagina kan per tabel de definities per kolom weergegeven worden. Dit zijn de definities die eerder in het proces opgesteld zijn. Hiermee voorkom je onduidelijkheden.



Datakwaliteit definities

Kies tabel

All

Gebruik de filter bovenaan de pagina om de Data rules per kolom te weergeven.

dim_leverancierkenmerken is enkel getoetst op administratie B110.

tabel	kolom	Gewicht	Data Rules
d_leverancier	Administratie	2	B800 series worden geen leveranciers ingelezen. B800 series is fout.
d_leverancier	Bestelmedium	2	De omschrijving moet gelijk zijn aan B110 omschrijving
d_leverancier	Betalingsconditie	2	0 is niet fout, leeg en dubbele waarden wel
d_leverancier	Betalingsconditie_crediteur	2	mag niet leeg gelaten worden, dubbel moeten hetzelfde zijn. 0 is niet fout.
d_leverancier	Betalingstermijn_crediteur_dagen	2	Verschillende 0 waarden zijn incorrect, leeg is fout en mag geen dubbele waarden bevatten.
d_leverancier	Ean_adres	0	
d_leverancier	Emailadres	2	Als de bestelmedium via e-mail is, moet er een emailadres zijn. Andere gevallen is leeg niet fout
d_leverancier	Inkoopconditie	1	Croon raamcontract is niet fout en overeenkomst W&D is niet fout. Andere dubbel wel. 0 en 99 zijn fout
d_leverancier	Kernleverancier	3	mag ja/nee zijn
d_leverancier	Leverancier_skey	3	
d_leverancier	Leveranciers nr	2	B929 heeft andere leveranciersnummers voor dezelfde leveranciers dan alle andere Administraties. Advies geven
d_leverancier	Leveringsconditie	1	0 waarden is fout, in dat geval wordt het veld leeg gelaten
d_leverancier	Naam	2	Naam mag niet leeg gelaten worden
d_leverancier	Status	2	Is verplicht en mag niet leeggelaten worden
d_leverancier	Wka_leverancier	1	Mag niet leeg zijn
d_leverancierkenmerken	Admin_code	1	Enkel B110 en B900 series bevatten
d_leverancierkenmerken	Datum geldig tot en met	2	1901 is fout. Maximaal 5 jaar in de toekomst, alles daarboven is fout.
Total		45	

Figuur 20 DQ definitie pagina

6 Conclusie

Het doel van dit onderzoek is een eindproduct opleveren die antwoord geeft op de vraag: 'Welke maatregelen dienen gemaakt te worden om de datakwaliteit te waarborgen wanneer de data toegankelijkheid verhoogd wordt.' Hiervoor is een kwalitatief en kwantitatief onderzoek naar gedaan, met als eindresultaat een proof of concept.

Voordat een nieuw bedrijfsproces ontwikkeld kan worden, moet eerst de huidige situatie duidelijk omschreven zijn. Hierin worden de knelpunten vastgesteld en wordt er een gewenste situatie of doelstelling geformuleerd.

Als een datamanagement tool binnen een organisatie wordt geïmplementeerd is het ook van belang om vast te leggen waar de tool voor gebruikt gaat worden en wie de tool gaan gebruiken om aan de wensen en eisen van de opdrachtgever en de stakeholders te voldoen. Er dient rekening gehouden te worden met de verschillende niveaus, van strategisch tot aan operationeel niveau. De rol van data dient duidelijk te zijn binnen de organisatie.

Het opstellen van een Data Governance council op strategisch niveau en data stewards aanstellen op tactisch niveau is een eerste stap naar datamanagement. Een Data governance council helpt de organisatie gegevens efficiënter te beheren. Het moet leiden tot consistente definities, vastgelegde procedures en afspraken die ervoor zorgen dat de data betrouwbaar, beschikbaar en consistent is. De data stewards leggen de definities, procedures en kwaliteitseisen vast. Deze worden uitgevoerd op operationeel niveau. Dit is onderdeel van een Data Governance project. Data governance zorgt voor de regie op en de sturing van de datamanagement activiteiten zodat het consistent wordt doorgevoerd binnen de hele organisatie.

Een datakwaliteit tool kan op verschillende manieren ontwikkeld worden. Om een juiste tool te kiezen is er onderzoek gedaan. De tool(s) word(en) gekozen op basis van de doeleinden waar ze gebruikt voor worden. De functionaliteiten en de doeleinden waar de tool voor wordt ingezet, worden tegenover elkaar afgezet. De resultaten tonen dat SQL Server meer mogelijkheden heeft om functionaliteiten te ontwikkelen. Op het gebied van performance scoort Power BI laag in vergelijking tot SQL Server, maar scoort hoog met data-analyse en visualisatie.

De datakwaliteit tool die is ontwikkeld voor dit onderzoek is schaalbaar en kan over de hele organisatie geïmplementeerd worden. Data problemen kunnen hiermee geïdentificeerd en geprioriteerd worden. Samen met business professionals en IT-professionals kan deze oplossing op alle afdelingen van de organisatie worden gebruikt. Uit de resultaten blijkt dat het volgens best practices gedaan dient te worden met een Data Governance Council, die de definities, procedures en regels op een hoog niveau vastleggen voor elke afdeling.

Er zijn meerdere bevindingen gedaan tijdens het onderzoek omtrent datakwaliteit. De gevolgen van de bevindingen hebben invloed op het inkoopdashboard en de bedrijfsprocessen. Door deze problemen zijn sommige delen van het dashboard niet betrouwbaar en hierdoor niet te gebruiken voor de business.

De visualisatie van het datakwaliteit dashboard maakt het mogelijk voor de business gebruikers om inzicht te krijgen over de mate van de datakwaliteit. De tool voldoet aan alle wensen en eisen van de opdrachtgever en de stakeholders.

6.1 Discussie

Uit de resultaten van de workshop met de inkopers is gebleken dat deze duidelijk weten te omschrijven wat de knelpunten en doelstellingen zijn. De inkopers willen de informatie van de leveranciers analyseren in het dashboard, zonder zich zorgen te hoeven maken of de informatie die weergegeven wordt wel of niet klopt met de werkelijkheid. Met die informatie onderhandelen de inkopers met leveranciers en kunnen op basis van de data de best mogelijke (bonus)afspraken maken.

Uit de stakeholdersanalyse is gebleken dat het management enthousiast is over de datakwaliteit tool. Een datakwaliteit oplossing welke binnen de hele organisatie wordt gebruikt. Het is noodzaak volgens de stuurgroep om inzicht te hebben in de kwaliteit van de data. Hoge kwaliteit van data heeft een positieve invloed op het maken van beslissingen op basis van data.

Het proof of concept is ontwikkeld met een raamwerk en is gevalideerd door de opdrachtgever, de stakeholders, het BI team en het management van inkoop. Alle partijen zijn tevreden met de resultaten en plannen om het in productie te nemen. Indien een inkoper een analyse doet op leveranciersgegevens en het datakwaliteit dashboard toont aan dat de leverancierstabellen met 99% correct zijn, kan de inkoper met veel zekerheid concluderen dat de data overeenstemt met de werkelijkheid. Op basis van de verwachtingen en de resultaten kan gesteld worden dat de eisen en wensen van het onderzoek zijn voldaan.

Hierbij moet wel rekening gehouden worden dat er veel datakwaliteit tools in omloop zijn. Dit product heeft zich enkel op de data controle gefocust van een kleine hoeveelheid data. Er kan geen uitspraak gedaan worden of het grootschalig implementeren van deze tool probleemloos zal verlopen. Meer functionaliteiten toevoegen kan veel tijd en geld kosten. Er zal een keuze gemaakt moeten worden of Croonwolder&dros zelf de tool verder ontwikkelen, of dat er een tool wordt gekocht. Het management heeft tijdens de doorlooptijd van het project actief gewerkt aan definities en is van plan om een Data Governance project op te zetten.

De inhoud van de oplossing bestaat uit foutieve data identificatie. Tijdens het onderzoek zijn meer problemen ontdekt die niet binnen de scope van dit onderzoek valt. Data Matching en Data Clustering zijn interessante functionaliteiten die de kwaliteit van de data significant kunnen verhogen. Hier zal een vervolgonderzoek naar gedaan moeten worden.

6.2 Aanbevelingen

De organisatie kan de oplossing in productie doorvoeren en op meerdere tabellen toepassen. Waarom niet een volledige tool aanschaffen en implementeren? Croonwolter&dros heeft tot op heden geen datamanagement. Het is nog niet duidelijk op welke schaal datamanagement opgezet gaat worden en of er een budget beschikbaar voor is. Het belang en invloed op de business van data management is nog niet bekend binnen de hele organisatie. Datakwaliteit management is een continue proces. De oplossing van dit onderzoek heeft meerdere datakwaliteit issues geïdentificeerd en is een goed uitgangspunt om vanuit door te ontwikkelen. Door de activiteiten te volgen (Zie “Hoe wordt de datakwaliteit getoetst?”) kan er gestructureerd doorontwikkeld worden om een meetinstrument op maat te maken.

Ik adviseer de organisatie een Data Governance Council op te zetten. De Council heeft één lid uit elke afdeling. De Council legt definities, procedures, strategie, beleid, normen & kwaliteit en naleving omtrent data management vast. Al deze gebieden worden organisatie breed toegepast. Zoals eerder omschreven is data management een continue proces. Een Data Governance Council verandert niet in één keer alle werkwijzen of procedures, naar aanleiding van de prioriteiten van de organisatie kunnen deze worden ingevoerd.

Als een Data Governance Council wordt opgesteld, is het noodzakelijk om data stewards aan te stellen om de gemaakte definities, procedures, strategie, beleid, normen & kwaliteit en naleving uit te voeren op tactisch niveau. Het is de verantwoordelijkheid van de data stewards dat de gemaakte afspraken van de Data Governance Council ook daadwerkelijk worden nageleefd op operationeel niveau.

7 Competenties

7.1 A-competenties

De A-competenties zijn afgeleid van algemene HBO A-competenties, die uitspraken doen over de houding en vaardigheden van alle HBO studenten. Hieronder per competentie een toelichting waarom de competentie van belang is voor het afstuderen van een Informatica student op de Hogeschool Leiden en waar de competentie vandaan komt.

7.1.1 Onderzoek

De deelvragen zijn op een manier geformuleerd dat, wanneer de lezer de informatie van de deelvragen heeft opgedaan hij/zij de hoofdvraag kan beantwoorden. De vragen zijn zo SMART als mogelijk gedefinieerd (Zie Onderzoeksvragen).

De methoden die gebruikt zijn gedurende het project staan beschreven in hoofdstuk 6 Methodologie. Voor iedere deelvraag is er een methode toegepast om deze te beantwoorden, bijvoorbeeld bij vergelijkend onderzoek bij hoofdstuk **Fout! Verwijzingsbron niet gevonden..** De juiste methoden zijn toegepast op de correcte vraagstellingen.

De bronnen die via internet of boeken zijn gebruikt staan in de Bibliografie en worden gebruikt als onderbouwing bij het beantwoorden van vraagstukken. Doormiddel van het organiseren van een workshop is er met post-its gewerkt om de requirements van de stakeholders te achterhalen. Deze workshop is toegevoegd als bijlage (*Bijlage 3 Workshop stakeholders presentatie*).

De resultaten van deelvragen worden altijd onderbouwt door een visualisatie en/of een onderzoeksmethode. Dit is terug te zien in het hoofdstuk Onderzoeksresultaten. De analyses zijn met de data analist, de data steward en/of de BI manager gevalideerd, om er zeker van te zijn dat de functionaliteiten correct zijn ontwikkeld en de wensen van de stakeholders worden gevolgd.

Per deelvraag is er een conclusie getrokken. In (hoofdstuk 6 Conclusie) is een conclusie getrokken op basis van de resultaten van de deelvragen en is er een advies geschreven om het product verder in de organisatie te implementeren.

7.1.2 Leren leren

Het ontwikkelen van een datakwaliteit tool is uitgevoerd volgens het DAMA DMBOK2 raamwerk. Het raamwerk is zelfstandig gekozen en goedgekeurd door de stagebegeleider. In overleg met de BI manager is de keuze gemaakt om enkel de domeinen van DAMA DMBOK2, die op dit moment actief een rol spelen binnen de organisatie verder op in te zoomen. Vervolgens wordt er nog dieper op datakwaliteit management ingezoomd en wordt hier een product voor ontwikkeld. De reden om maar drie domeinen te kiezen is omdat anders te veel tijd verloren gaat aan activiteiten die niet of minder relevant zijn aan het project.

Tijdens de opdracht heb ik verschillende tools en methoden geprobeerd om de opdracht te maken (Zie Power BI versus SQL Server). Sommige tools waren minder goed geschikt en andere waren beter geschikt voor de doeleinden. Om de datakwaliteit te meten en te kunnen beheren, zijn er analyses met Power BI uitgevoerd. Power BI had een aantal nadelen op het gebied van het toepassen van de business op datakwaliteit aspecten.

De datakwaliteit is in SQL Server uitgevoerd en vervolgens in Power BI geladen om te visualiseren. De business rules zijn in SQL verwerkt en de visualisaties en berekeningen zijn vervolgens in Power BI geïmplementeerd. De performance van het verwerken van de data gaat significant achteruit als er veel aanpassingen op tekstvelden wordt gedaan binnen Power BI. Dit is nog een reden geweest om zoveel mogelijk activiteiten en functionaliteiten in SQL Server te verwerken.

Met het ontwikkelen van de tool zijn er een aantal bevindingen geweest in de data waar de medewerkers van de IT-afdeling niet vanaf wisten. Deze bevindingen worden nu naar de ERP-systeem leverancier gecommuniceerd voor een oplossing.

7.1.3 Professioneel werken

Het probleem wordt beschreven in [Bijlage 14 Plan van Aanpak](#) en in het hoofdstuk Inleiding. Er is altijd volgens een methode gewerkt. De medewerkers van de IT-afdeling zijn up-to-date gebleven gedurende het project. Er zijn tussentijdse evaluaties geweest met de IT-afdeling en met de inkoopafdeling. Gedurende het project is er via de e-mail en in-person contact geweest met de opdrachtgever. Aan het einde van iedere sprint/iteratie heeft een review moment plaatsgevonden met de opdrachtgever. Tijdens het tonen van het werk werd er feedback gegeven op de huidige status van het project en waar de organisatie naartoe wil met het product.

Het meetinstrument is zelfstandig opgesteld. Tijdens een bespreking van het meetinstrument heeft de BI-manager geadviseerd om enkel drie domeinen van het DAMA DMBOK te behandelen en vervolgens verder op datakwaliteit in te zoomen. Er zijn verschillende tools geprobeerd om de datakwaliteit op te stellen, te meten en om te beheren. Dit is overlegd met de BI-manager en de data analist van Croonwolter&dros.

7.1.4 Innovatie

De organisatie heeft voorafgaand aan het project nog geen datamanagement proces of tool geïmplementeerd. Hierdoor sta ik in de positie om invloed te hebben om verschillende opties toe te passen. Hoe het product gerealiseerd moet worden of hoe het eruit komt te zien was nog niet duidelijk. In het hoofdstuk “Hoe wordt de datakwaliteit getoetst?” is een datakwaliteit tool ontwikkeld met behulp van het functioneel kader DAMA DMBOK2. Er is tijd geïnvesteerd om de juiste software te gebruiken om het product schaalbaar en beheersbaar te maken. Een vereiste requirement is dat de business user met het product zelfstandig kan werken.

7.2 B-Competenties

De B-competenties zijn per deelvraag verdeeld. Iedere deelvraag staat onder een B-competentie met een korte uitleg. De competenties worden aangetoond in de onderzoeksresultaten.

7.2.1 Bedrijfsprocessen analyseren

5.1.1 Welke KPI's en requirements zijn gedefinieerd binnen de inkoopafdeling?

Dit is de eerste deelvraag die is uitgevoerd in het project. Doormiddel van een workshop met vijf inkopers en de opdrachtgever is het bedrijfsproces in kaart gebracht. De huidige situatie is vastgelegd, met de knelpunten. Vervolgens zijn de wensen, eisen en de doelstellingen geïdentificeerd. De knelpunten, wensen en doelstellingen zijn via post-its geprioriteerd met de MoSCoW methode ([Bijlage 3 Workshop stakeholders presentatie](#) & [Bijlage 1 Requirements en KPI's in DevOps](#)). Op basis van de workshop zijn de KPI's en de requirements gedefinieerd.

5.1.3 Wat doet de organisatie aan data management?

Voor het beantwoorden van deze deelvraag is een interview gehouden met de informele data steward van de inkoopafdeling. Hieruit is gebleken dat er geen gestructureerde manier van data management aanwezig is. Op basis van dit interview en de resultaten kan er herleid worden dat data kwaliteit issues met het oog worden geïdentificeerd. De organisatie heeft geen inzicht op de mate van de kwaliteit van de data. Mede door dat er geen data management aanwezig is, wist de organisatie niet goed wat ze precies wilde ontwikkelen.

5.1.8 Welke opmerkelijke datakwaliteit gerelateerde bevindingen zijn er gevonden?

Gedurende het project zijn er meerdere knelpunten geïdentificeerd en opmerkelijke bevindingen gedaan. De invloed van het probleem van de projectnummers die niet goed in Acto staan is groot. Hier is een kwalitatief onderzoek toegepast. Het Project deel van het inkoopdashboard voor de inkoopafdeling is niet betrouwbaar en onvolledig. Hierdoor kan het niet gebruikt worden. Om dit probleem te achterhalen is er met een business expert gesproken. Met de expert is er stapsgewijs naar het probleem gezocht en is er een conclusie getrokken.

Het andere probleem omtrent data matching, is ook een datakwaliteit probleem. Hier is een kwantitatief onderzoek naar gedaan. Er is een voorbeeld genomen in de deelvraag waarbij de verschillen zichtbaar zijn. Het belang en het risico van het probleem is beschreven. In SQL Server en Power BI zijn de bevindingen dubbel gecontroleerd en hierin is vastgesteld dat het probleem op grote schaal aanwezig is binnen de organisatie.

7.2.2 Infrastructuur analyseren

5.1.4 Welke invloed heeft de data migratie naar Microsoft Azure op het project?

Voor het beantwoorden van de invloed van de migratie naar Azure is er onderzoek gedaan naar de huidige situatie en de gewenste situatie. In het hoofdstuk worden de knelpunten beschreven en hoe deze opgelost kunnen worden met het migreren naar Azure. De nieuwe infrastructuur in Azure staat beschreven bij Figuur 6 ETL-proces Azure inclusief de voordelen die hiermee gepaard gaan. Er wordt verder naar de toekomstige situatie gekeken en daar wordt rekening mee gehouden.

5.1.7 Welke tools kunnen gebruikt worden voor datakwaliteit management?

In overleg met de data analist, database beheerder en de BI manager is er gekozen om de datakwaliteit tool met software te ontwikkelen dat gebruikt wordt binnen de organisatie. Er is een vergelijkend onderzoek toegepast tussen de mogelijke software programma's waarin de tool in gebouwd kan worden. Een belangrijke requirement van de tool is dat het te integreren is met de huidige bedrijfsprocessen. Daarnaast zijn verschillende mogelijkheden overwogen tussen het aanschaffen en zelf ontwikkelen van een tool. Als de organisatie ervoor kiest om zelf de tool verder te ontwikkelen, is het een must dat dit mogelijk is in de gekozen software.

7.2.3 Bedrijfsprocessen ontwerpen

5.1.5 Welke aspecten van datamanagement hebben een impact op dit project?

Deze deelvraag is uitgewerkt op basis van het theoretisch kader, ook wel het meetinstrument. Er is onderzoek gedaan naar drie domeinen van DAMA DMBOK2; Data Governance, Datakwaliteit- en Data Security management. De werkwijzen van het theoretisch kader zijn samengevoegd met de business. Er is op verschillende niveaus een Data Governance bedrijfsproces op hoog niveau beschreven en er wordt draagvlak gecreëerd. Op het gebied van Datakwaliteit worden voornamelijk richtlijnen beschreven. De reden dat ervoor gekozen is om datakwaliteit niet in dit hoofdstuk uitgebreid uit te schrijven, is omdat er een proof of concept van datakwaliteit management wordt ontwikkeld. Op het gebied van Data Security worden meerdere aandachtspunten benoemd en hier wordt een advies op gegeven.

5.1.6 Hoe wordt de datakwaliteit getoetst?

Deze deelvraag heeft inhoudelijk raakvlakken met “Bedrijfsprocessen analyseren” en “Infrastructuur analyseren”. In dit hoofdstuk staat beschreven hoe de datakwaliteit tool is opgezet en hoe de data wordt gemeten. De tool is ontwikkeld met behulp van het DAMA DMBOK2 raamwerk en volgt zes gestructureerde stappen om tot het eindresultaat te komen.

De eerste stap die gemaakt is om de tool te ontwikkelen, is door definities vast te leggen over datakwaliteit, de knelpunten en de huidige situatie. Vervolgens wordt de strategie bepaald waarop de datakwaliteit op wordt gemeten en beheert. Hierbij is er een keuze gemaakt voor een gestandaardiseerde formule waarmee dezelfde input altijd dezelfde output geeft. Met behulp van de KPI's en de requirements zijn de gewichten van de data velden geprioriteerd. Het gewicht geeft aan hoe belangrijk het data veld is voor de business. Naderhand wordt er het belang van data stewards beschreven, waardoor het draagvlak wordt vergroot en de stakeholders worden betrokken bij het proces. Uit eigen onderzoek zijn verschillende datakwaliteit tools vergeleken met elkaar op het gebied van performance, gebruiksvriendelijkheid en functionaliteiten om de beste keuze te maken. Verder is er een eerste meting gedaan die later is gevalideerd. De bevindingen worden geprioriteerd door de business professionals. Zij bepalen of de foutieve data wordt hersteld in het datawarehouse. De doelen van de verbeteringen van datakwaliteit worden door de data stewards of door het management bepaald.

5.1.6.6 Datakwaliteit modificaties ontwikkelen en implementeren

Tenslotte wordt in de laatste activiteit van de deelvraag “Hoe wordt de datakwaliteit getoetst?” de tool geïmplementeerd. In overleg met de data analyst is besloten om in SQL Server tabellen aan te maken, waarin de resultaten van de metingen worden opgeslagen. Na de implementatie van de tool heeft er nogmaals een validatie plaatsgevonden, om de kwaliteit van de tool te waarborgen.

5.1.9 Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?

Voordat het dashboard is gemaakt, zijn de wensen en eisen van de stakeholders vastgelegd. De look and feel van het dashboard moet overeenkomen met het inkoopdashboard. Voorafgaand is een schets gemaakt van het dashboard samen met de data steward. Hierin zijn de functionaliteiten besproken en de wensen van de visualisatie. Vervolgens is het dashboard gemaakt en is er feedback gegeven door de data steward en de opdrachtgever. Een definitie tabblad is een eis, zodat de gebruikers kunnen zien waar de data velden op zijn getoetst. Het dashboard kan gebruikt worden door business professionals, waardoor zij zelf verantwoordelijk worden voor de datakwaliteit van de afdeling.

8 Reflectie

Dit hoofdstuk beschrijft de reflectie en de wijzigingen die gedurende het project zijn doorgevoerd in het project.

8.1 Planning wijzigingen

Aan het begin van het project is de initiële planning gemaakt. Hierin is aangegeven dat het project binnen de gestelde tijd voltooid dient te zijn. Gedurende het project zijn er een aantal activiteiten gewijzigd in de planning. De requirements achterhalen en vervolgens Microsoft DevOps inrichten duurde langer dan gepland is. De requirements waren in eerste instantie niet goed te differentiëren van elkaar in de Product Backlog. Dit kwam omdat meerdere requirements dezelfde data nodig heeft voor een analyse, maar enkel door een andere filter toe te passen (project, divisie of leverancier) een andere requirement opgesteld werd. De requirements zijn vervolgens aangescherpt en onder drie verschillende domeinen geplaatst (Zie Tabel 6 Key performance indicatoren & [Bijlage 1](#)) en de filters zijn verwijderd uit de requirements en later als pagina filter toegevoegd in de rapportages. Alle planningswijzigingen zijn in DevOps aangepast en doorgevoerd.

Een onderzoek naar externe databronnen is achterwege gelaten. Hier is geen tijd voor geweest door problemen met het kiezen van de software waar de datakwaliteit tool mee wordt ontwikkeld. Om die reden is er geen Deliverable 7 “Externe data onderzoeksrapport”.

Er is geen individueel technisch ontwerp opgesteld. Dit is gedeeltelijk opgenomen in het hoofdstuk (Eerste datakwaliteit beoordeling). Er zijn twee nieuwe tabellen toegevoegd aan het huidige model. Deze staan in relatie tot elkaar met surrogate keys.

Het datakwaliteit rapport is verdeeld beschreven in het afstudeerverslag. Het hoofdstuk “Welke opmerkelijke datakwaliteit gerelateerde bevindingen zijn er gevonden?” & “Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?” behandelen de bevindingen die gedaan zijn omtrent datakwaliteit en het visualiseren van de huidige status van de datakwaliteit. Er is gekozen om geen apart datakwaliteit rapport te maken.

In het PvA staat dat er een auditrapport wordt opgeleverd. Hier is geen document voor gemaakt, maar de validatie is gedurende het proces uitgevoerd. Dit is in samenwerking gedaan met de BI-manager, de data analist en de data steward. In het hoofdstuk “Datakwaliteit modificaties ontwikkelen en implementeren” is er een test uitgevoerd op de correctheid van de tool.

Het technisch ontwerp is verwerkt in de deelvragen “Welke invloed heeft de data migratie naar Microsoft Azure op het project?” & “Datakwaliteit modificaties ontwikkelen en implementeren”. Er is voor gekozen om dit in het afstudeerverslag te plaatsen en geen apart document voor te schrijven.

Er zijn gedurende het project een aantal deelvragen aangepast. De hoofdvraag is meer leesbaar en begrijpelijk gemaakt. De deelvragen zijn aangepast om beter aan te sluiten aan het onderzoek en activiteiten om tot het eindresultaat te komen.

De Risicoanalyse in het PvA is niet meer relevant geweest voor de scope van mijn project. Gedurende het project is er besloten dat ik niet ad-hoc de datakwaliteit problemen aan ging pakken, maar een gestructureerde tool ontwikkelen om de datakwaliteit te meten. Met deze tool zullen foutieve data en verkeerde relaties tussen tabellen geïdentificeerd worden. Het voordeel hiervan is dat niet alleen de zichtbare problemen worden geïdentificeerd, maar ook de achterliggende problemen omdat alle velden worden meegenomen in de meting.

Op 23 mei is er een demo presentatie geweest voor een datakwaliteit tool. Dit product DataMatch Enterprises van DataLadder zou een optie kunnen zijn om aan te schaffen, om datakwaliteit te beheren en te verbeteren. Dit onderzoek is geen deel van de scope geweest en is later toegevoegd. Er was geen tijd over om een business case te creëren.

Overige bijlagen

Als externe bijlage zijn de volgende documenten toegevoegd, om het afstudeerdossier compleet te maken:

- Interview feedback
- Afstudeervoorstel
- Bijlage A Studieplanning
- Bijlage D Adviesformulier
- Portfolio ISLH
- Plan van Aanpak
- Meetinstrument

8.2 Evaluatie

Wat was de situatie?

Het afstudeertraject is van start gegaan vanaf 11 februari 2019 bij Croonwolter&dros. Samen met een mede-afstudeerder Sike van Dongen hebben we beide een deel van de opdracht uitgevoerd in opdracht van de inkoopafdeling. De afdeling waar wij bij in dienst waren was de BI-afdeling. De betrokkenen bij het project waren de BI-manager en voornamelijk de data analist en databasebeheerder. Croonwolter&dros is in februari begonnen met migreren naar Azure en hebben tegelijkertijd nieuwe data marts aan laten leveren door Acto. Acto is de leverancier van het CRM-systeem van Croonwolter&dros. De inkoopafdeling wilde inzicht krijgen in de inkoopdata gecombineerd met de crediteuren data, zodat er een compleet overzicht is van de leverancier. Niet enkel de bestelorders, maar ook de facturen en of leveranciers en Croonwolter&dros zich aan de betaal- en levervoorwaarden voldoen. Er was nog geen datamanagement proces aanwezig en dit houdt in dat de organisatie geen inzicht heeft op de kwaliteit van de data, noch dat er een gestructureerde manier is vastgelegd om data gerelateerde issues op te lossen. Definities zijn per afdeling vastgelegd en de afdelingen opereren vanuit hun eigen oogpunt, zonder kennis te hebben hoe andere afdelingen performen.

Wat was mijn taak?

Mijn rol in dit project was om een datamanagement proof of concept te ontwikkelen. Deze proof of concept moet voldoen aan bepaalde requirements, zodat het over de hele organisatie geïmplementeerd kan worden. De tool heeft als doel om een gestandaardiseerde score terug te geven, die een indicatie geeft over de status van datakwaliteit van een data asset. De Business Professionals, in dit geval de inkopers moeten met deze tool zelfstandig de datakwaliteit te kunnen beheren en verbeteren. De verwachtingen waren dat ik alle data marts van de inkoopdata zou implementeren in de tool. Mijn verwachtingen waren dat ik de tool, zoals hierboven is beschreven te kunnen implementeren en volledig automatisch integreer in het systeem. Mijn doel van het afstudeertraject is om zelfstandig een datakwaliteit meetinstrument te kunnen integreren in een organisatie, die aansluit aan de behoefte van de organisatie. Een anders doel was om alle domeinen van het DAMA DMBOK uit te voeren tijdens mijn afstudeertraject, om een compleet datamanagement proces bij Croonwolter&dros te integreren.

Hoe heb ik het aangepakt en waarom?

De eerste weken heb ik veel meetings gehad met uitleg over de bedrijfsprocessen en de infrastructuur van het datawarehouse en de proces flow van data en activiteiten. Vervolgens zijn de requirements van de stakeholders en de opdrachtgever opgesteld. De requirements waren een rode lijn in het traject. Om het traject succesvol af te ronden, is een van de doelstellingen om de stakeholders en de opdrachtgever tevreden te houden. Vervolgens heb ik het DAMA DMBOK2 raamwerk als functioneel kader/theoretisch kader gebruikt voor mijn onderzoek. Ik heb een theoretisch kader ontwikkeld die in mijn ogen belangrijk waren voor de organisatie. Per domein heb ik de relevante onderdelen gedocumenteerd en toegepast op de business van Croonwolter&dros. Drie van de domeinen zijn voor verder onderzoek geselecteerd. Waarom deze drie? Op het moment van het afstudeertraject was de organisatie actief bezig met het wijzigen van Data Security en speelde er bij het management Data Governance vraagstukken, zoals definities definiëren, KPI targets en het data management. Datakwaliteit management stond centraal in mijn onderzoek, ook wel het doel van mijn afstuderen. Zelfstandig heb ik gewerkt naar een eindproduct, omdat er via de SCRUM-methode gewerkt werd, hebben we iedere twee dagen (dinsdag en donderdag) daily stand ups gehouden. Tijdens deze korte meetings werd besproken wat er gedaan was de voorgaande dagen en waar we de komende dagen aan gingen werken. De knelpunten werden benoemd als die er waren, maar hier werd tijdens de daily stand up niet verder op ingegaan. De tussenresultaten werden wekelijks besproken met de BI-manager en de opdrachtgever. De data analist heeft mij ondersteuning geboden waar nodig.

Heeft het gewerkt?

Het DAMA DMBOK2 is een enorme hulp geweest met het opstellen van een gestructureerde manier van het opzetten van een datamanagement tool. Het bevat de bouwstenen waarmee naar het eindproduct gebouwd kan worden. De theorie moet wel toegepast worden op de business van de organisatie. De SCRUM-methode was een goede methode, omdat er op een gestructureerde manier contact is tussen de projectleden, de opdrachtgever en de stakeholders.

Wat heb ik ervan geleerd?

Ik ben altijd erg kritisch naar mijzelf toe. De verwachtingen van mijzelf zijn hoog. Halverwege het traject was ik niet tevreden met het product en heb ik in de weekenden verder gewerkt. Het eindresultaat is in mijn ogen op een klein gedeelte van het geheel uitgevoerd. Dit is jammer, maar het eindproduct bevat wel alle inhoudelijke componenten van de tool. Het toepassen van de tool op meer datasets is het herhalen van dezelfde activiteiten en technieken. Ik ben tevreden met het eindresultaat, de tool is schaalbaar, geautomatiseerd en geeft een gestandaardiseerde eindscore aan per tabel/kolom/filter. De inkoopafdeling kan zelfstandig de datakwaliteit beheren en verhogen. Dat is uiteindelijk wat het doel was vanaf het begin van het afstudeertraject.

Ik heb geleerd om procesmatig en meer gestructureerd te werken. Communicatie is een van de meest essentiële onderdelen in projecten. De volgende keer zou ik meer contact opnemen met de stakeholders en de opdrachtgever. Uiteindelijk ontwikkel je het product voor hun en ze weten wat ze willen. Ik bezit nu de kennis om een datakwaliteit management tool te kunnen opzetten en initiëren. Ik beheers de SCRUM-werkmethode en kan deze toepassen in andere situaties.

Tenslotte wil ik de collega's bij Croonwolter&dros bedanken voor de ondersteuning en gezellige tijd. Ze stonden altijd klaar voor mij wanneer ik vragen had of iets niet volledig begreep. Zonder de goede begeleiding van Rob Kabel en Albert Huttenga heb ik niet tot dit eindproduct kunnen komen.

9 Bibliografie

- Acto. (2019, 05 27). *Acto*. Opgehaald van Acto: <https://www.acto.nl>
- arbocatalogus, d. (2019, 02 14). *arbocatalogus*. Opgehaald van deArbocatalogus: <https://www.dearbocatalogus.nl/handreiking-intern-transport-en-logistiek/52-beoordelingstool-2-risicobeoordeling-van-fine-en-kinney>
- Brautigam, I. (2019, 02 19). *Scrumguide*. Opgehaald van Scrumguide: <https://scrumguide.nl/sprint-review/>
- DAMA. (2017). *DAMA DMBOK2*. -: DAMA.
- DataLadder. (2019, 05 25). *DataLadder*. Opgehaald van DataLadder: <https://www.Dataladder.com>
- ERP. (2019, 02 21). *ERPdefinitions*. Opgehaald van ERPdefinitions: <https://www.erpsystemen.nl/wat-is-erp>
- Fred Conijn, J. d. (2018, 01 01). *cmweb*. Opgehaald van cmweb: <https://cmweb.nl/2018/01/datamanagement-kwaliteit-gegevens-bepaalt-kwaliteit-informatie/>
- Jonker, R. (2019, 02 28). *Compact*. Opgehaald van Compact: <https://www.compact.nl/articles/datakwaliteitsonderzoek/>
- Kimball. (2019, 05 20). *KimballGRoup*. Opgehaald van KimballGroup: <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/>
- Leansixsigma. (2019, 02 06). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/overzicht-stakeholders>
- Leansixsigma. (2019, 05 07). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/wat-zijn-key-performance-indactors>
- Microsoft. (2019, 4 20). *azure.microsoft*. Opgehaald van azure.microsoft: <https://azure.microsoft.com/nl-nl/overview/what-is-devops/>
- Motion10. (2019, 02 06). *Motion10*. Opgehaald van Motion10: <https://www.motion10.nl/>
- Oreilly. (2019, 03 14). *Oreilly*. Opgehaald van Oreilly: <https://www.oreilly.com/topics/oreilly-learning>
- Redactie. (2012, 10 26). *Business analytics*. Opgehaald van Business analytics: <https://www.business-analytics.biz/blogs/de-waarde-van-data-governance-en-het-zichtbaar-maken-ervan/>
- redactie, C. (2019, maart 22). *Computerworld*. Opgehaald van Computerworld: <https://computerworld.nl/business-intelligence/61874-wat-is-datawarehousing>
- Swaen, B. (2019, 04 20). *Scribbr*. Opgehaald van Scribbr: <https://www.scribbr.nl/onderzoeksmethoden/kwalitatief-vs-kwantitatief-onderzoek/>
- Tozzi, C. (2019, 02 28). *Syncsort*. Opgehaald van Syncsort: <https://blog.syncsort.com/2018/02/data-quality/how-to-measure-data-quality-7-metrics/>
- University, S. (2019, 05 09). *Stanford*. Opgehaald van Stanford: <https://web.stanford.edu/dept/pres-provost/cgi-bin/dg/wordpress/wp-content/uploads/2012/09/DG-Data-Definition-Best-Practices.docx>



- Verheul, M. (2019, 05 16). *Agilescrumgroup*. Opgehaald van Agilescrumgroup:
<https://agilescrumgroup.nl/welke-rollen-zijn-er-binnen-scrum/>
- Wethinknext. (2019, 02 07). *wethinknext*. Opgehaald van wethinknext:
<https://wethinknext.com/stakeholder-mapping/>
- Wikipedia. (2019, 03 07). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Databeveiliging>
- Wikipedia. (2019, maart 19). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Data-architectuur>

10 Bijlagen

10.1 Bijlage 1 Requirements en KPI's in DevOps

De requirements zijn in Excel geformuleerd en geprioriteerd. Zie onderstaande afbeelding.

Requirements/ User Story				
Index N°	Als <gebruiker/rol>	will ik <functie/actie>	zodat ik <doel/waarde>	Prioriteits niveau
R1	Divisie Inkoop	de omzet per project kunnen analyseren per tijd dimensie	beslissingen kan maken op basis van de huidige omzet per project.	m
R2	Divisie Inkoop	de omzet per Derden binnen een VOF kunnen inzien	de bijdrage kan toetsen van CWD binnen een VOF.	w
R3	Divisie Inkoop	de omzet per productgroep		m
R4	Divisie Inkoop	de spend per reeds geleverd/uitgevoerd kunnen inzien van de bestelorders	de huidige status van de product flow kan registreren	s
R5	Divisie Inkoop	De omzet kunnen zien van bestelorders die niet via een purchase order gemaakt zijn	hier actie in kan ondernemen om het aantal non purchase orders te minimaliseren	kpi
R6	Category Manager	kunnen inzien welke leverancier een raamcontract heeft	hier rekening mee kan houden bij het kiezen van leverancier	m
R7	Category Manager	per leverancier de status (open/dicht) kunnen inzien	niet tijd verlies om een "dichtstaande" leverancier te contacteren	c
R8	Category Manager	per leverancier de kenmerken kunnen inzien	efficiënt en snel kan zoeken naar de juiste leverancier.	c
R9	Category Manager	per voorkeur/leverancier de productgroep en divisie kunnen inzien	geen tijd verlies aan het zoeken naar de voorkeur leveranciers per productgroep	m
R10	Project Inkoop	de omzet van voorkeur/leveranciers kunnen aflezen	kennis heb van de waarden per voorkeursleverancier	m
R11	Project Inkoop	de omzet van een leverancier kunnen inzien per divisie	per divisie de status kan controleren	kpi m
R12	Project Inkoop	de omzet van een leverancier kunnen inzien per regio	per regio de status kan controleren	m
R14	Inkoop Informatieanalist	kunnen inzien wat de omzet is per raamcontract	kan aflezen welk raamcontract de meeste omzet genereren	m
R15	Inkoop Informatieanalist	kunnen inzien wat de omzet per projectcontract	de verschillen per projectomzet kan achterhalen	w
R16	Inkoop Informatieanalist	de leveranciers categoriseren per leverpakket	sneller en efficiënter een leverancier kan vinden voor mijn doeleinden.	w
R18	Divisie Inkoop	kunnen inzien wat de status report is per project	een overzicht heb per project in welke fase het zich bevindt en welke handelingen er verricht moeten worden.	c
R19	Divisie Inkoop	rubrieken van 'derden'	???	w
R20	Divisie Inkoop	rubrieken van VKL	???	w
R21	Divisie Inkoop	een overzicht hebben van behaalde onderhandelingen	kan inzien welke onderhandelingen er in voorgaande afspraken zijn gemaakt.	w
R22	Divisie Inkoop	een overzicht hebben op de inkoop kraaijic analyse	in een oogopslag kan zien ...???	s
R23	Category Manager	de topdown omzet kunnen inzien van de leverancier vanuit de productgroep en subgroep	?? Wat voor activiteiten wil je uitvoeren op basis van deze informatie	m r3
R24	Category Manager	kunnen zoeken naar ik welke producten vandaan haal	snel de best passende leverancier kan vinden voor mijn doeleinden.	w
R25	Project Inkoop	inzicht hebben op de onderhandelings resultaten per order (grote orders, kleine orders)	kan inzien hoeveel rendement er behaald is per order.	c
R26	Project Inkoop	een overzicht hebben van waar wat is in te kopen (aanvulling op VKL lijst)	sneller en efficiënter een leverancier kan vinden voor mijn doeleinden.	w
R27	Project Inkoop	de omzet per VKL kunnen inzien	kennis heb van de waarden per voorkeursleverancier	m
R28	Divisie Inkoop	de omzet per kostenplaats kunnen inzien (in structuur ind. VOF)	?? Wat voor activiteiten wil je uitvoeren op basis van deze informatie	dim
R29	Project Inkoop	een overzicht hebben van de aantal projectcontracten per project	inzicht heb in hoeveel projectcontracten per project gebruikt worden.	w
R30	Project Inkoop	kunnen inzien aan welke projecten een leverancier levert	?? Wat voor activiteiten wil je uitvoeren op basis van deze informatie	m
R31	Divisie Inkoop	inzicht hebben op aantal projecten binnen een regio	een indicatie heb van de bezetting per regio.	dim
R32	Project Inkoop, Divisie Inkoop	de bonusafspraken per raamcontract kunnen inzien	hier op kan sturen met het inkopen bij raamcontract-leveranciers.	s
R33	Divisie Inkoop	de status van mijn werkbegroting kunnen zien	rekening kan houden met mijn te besteden begroting voor de aankomende periode.	w
R34	Divisie Inkoop	inzicht hebben op de kenmerken van de raamcontracten	de looptijd en specificatie van de actuele raamcontracten kunnen inzien	m
R35	Project Inkoop, Divisie Inkoop	status van de geleverde bestelorder per leverancier kunnen zien	?? Wat voor activiteiten wil je uitvoeren op basis van deze informatie	s
R36	Project Inkoop, Divisie Inkoop	kunnen inzien of de leverancier de bestelorder optijd levert	De lever-performance wordt meegenomen in het maken van een besluit tussen leveranciers.	m
R37	Category Manager	een overzicht hebben van de spend per inkoopcategorie	?? Wat voor activiteiten wil je uitvoeren op basis van deze informatie	c

Figuur 21 requirements Excel

Index N°	Prioriteits niveau	Dimensie	KPI	Target	Meting	Opmerkingen	Acto data	Niet Acto data
K1	M	Kostenbesparing	Percentage van het inkopen via raamcontracten (Aantal + Bedragen)		per maand/ kwartaal		TBI160	
K2	M	Kostenbesparing	De huidige status van inkopen bij leveranciers vs. De bonus targets		per maand/ kwartaal			Bonus (Excel)
K3	S	Kwaliteit	Percentage leverperformance per leverancier (gewenste leverdatum vs. daadwerkelijke leverdatum)		per maand/ kwartaal		Factuurdata	
K4	S	Kwaliteit	Percentage productkwaliteitscore per leverancier (volledig/niet volledig)		per maand/ kwartaal			
K5	M	Management	Aantal openstaande facturen per leverancier		per maand/ kwartaal			
K6	W	Management	Omzet percentage van een VOF (waar CWD penvoerder van is)		per maand/ kwartaal			
K7	M	Management	Omzet percentage van inkopen via VKL vs. Niet VKL		per maand/ kwartaal			VKL lijst (Excel)
K8	W	Management	Percentage/Aantal Status werkbegroting (Besteden vs not te besteden)		per maand/ kwartaal		Begroting is nog niet beschikbaar	
K9	C	Kostenbesparing	Omzet Percentage met PO		per maand/ kwartaal			
K10	S	Management	Percentage aantal leveranciers die leveren (vs. Totale leveranciers)		per maand/ kwartaal			

Figuur 22 KPI's Excel

Het Excel sheet is uitgewerkt in Microsoft DevOps onder "User Requirements". Zie onderstaande afbeelding.

De Data Driven CWD Inkoop Team									
+ New Work Item View as Board Column Options									
Order	Work Item Type	ID	Title	State	Priority	Effort	Busin...	Value Area	
1	Epic	504	> Issue lijst	New	2			Business	
2	Epic	467	> Data Management	New	2			Business	
3	Epic	446	> User Requirements	New	2			Business	
	Feature	374	> Leverancier Performance & Bewaking	Resolved	1			Business	
	Feature	449	> Omzet Beheer & Control	Active	1			Business	
	Feature	450	> Project Inkoop Beheer	New	2			Business	
	Feature	380	> Raamcontracten Algemeen	New	2			Business	
	Feature	451	> Purchase Order Beheer	New	2			Business	
	Feature	447	> Begroting Beheer	New	4			Business	
	Feature	448	> Overige requirements	New	2			Business	
4	Epic	415	> Non-user requirements	New	2			Business	

Figuur 23 DevOps product backlog

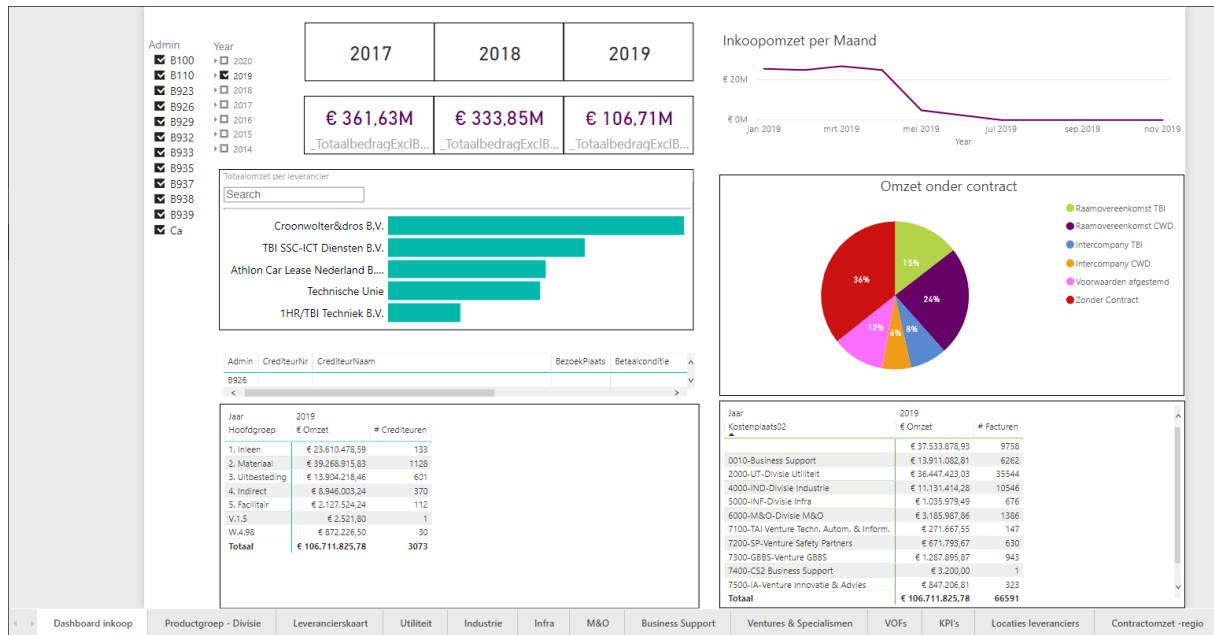
In onderstaande afbeelding zijn de user stories die onder het onderdeel Leverancier Performance & Bewaking zijn geplaatst.

Order	Work Item Type	ID	Title	State	Priority	Effort	Busin...	Value Area	Tags
2	Epic	467	> Data Management	New	2			Business	
3	Epic	446	> User Requirements	New	2			Business	
+	Feature	374	> Leverancier Performance & Bewaking	Resolved	1			Business	
	User Story	367	De huidige status van inkopen bij leveranciers vs. de bonus targets	New	4			Business	bonus KPI leverancier
	User Story	375	De leverperformance in % per leverancier kunnen inzien	New	4			Business	In DWH? KPI leverancier
	User Story	378	De productkwaliteitscore in % per leverancier kunnen inzien	New	4			Business	KPI kwaliteit leverancier
	User Story	381	Het aantal openstaande (nog niet betaalde) facturen per leverancier kunnen inzien	Closed	2			Business	factuur KPI leverancier
	User Story	385	De aantal leveranciers die leveren vs. totaal aantal leveranciers in % kunnen inzien	New	2			Business	KPI leverancier management
	User Story	370	De leveranciers die een raamcontract hebben kunnen inzien	Closed	1			Business	R14 R6 requirements
	User Story	377	De bestelorder(s) die op tijd opgeleverd zijn per leverancier kunnen inzien	Resolved	4			Business	R36 requirements
	User Story	376	De leverancier status (geblokkeerd/vrij voor gebruiker) per leverancier kunnen inzien	Resolved	3			Business	R7 requirements
	User Story	379	De nakoming van raamcontractvoorwaarden per leverancier kunnen inzien	New	2			Business	leverancier PurchaseOrder
	User Story	391	Het aantal Purchase Orders per leverancier kunnen inzien.	New	1			Business	leverancier PurchaseOrder
	User Story	395	De leverancier kenmerken per leverancier kunnen inzien	New	3			Business	kenmerken leverancier
	User Story	397	De leveranciers categoriseren per leverpakket	Closed	4			Business	leverancier leverpakket
	User Story	464	> Het definiëren van de 'Leverancier Performance & Bewaking' feature design	Closed	2			Business	Design
	User Story	493	> 'Leverancier Performance & Bewaking' Power BI Design & Build	New	2			Business	

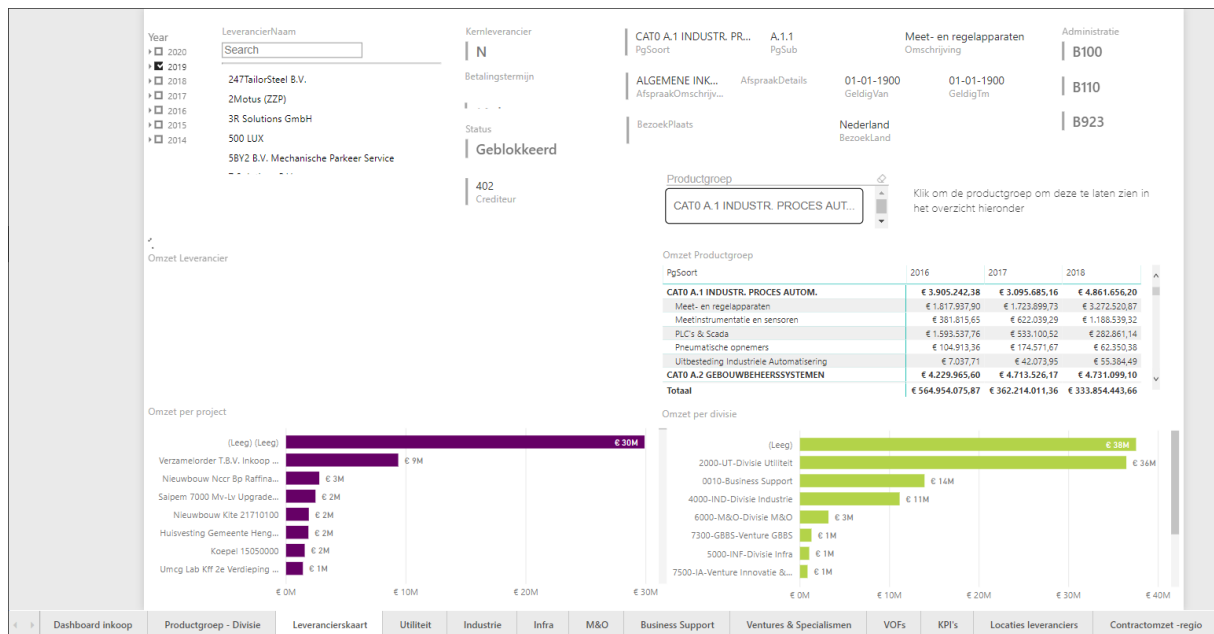
Figuur 24 Dev ops user stories

10.2 Bijlage 2 Oude dashboard inkoopafdeling

Onderstaande figuren weergeven de huidige rapportages die gebruikt worden door de inkoopers.



Figuur 25 Startpagina



Figuur 26 Leverancierskaart

10.3 Bijlage 3 Workshop stakeholders presentatie



Data Driven inkoop

15 minuten

Inkoopafdeling

Inhoud

3



Afstudeerproject

4



- ✓ **Inkoopproces**
Optimaliseren van het inkoopproces, om het werk makkelijker te maken voor de inkopers.
- ✓ **Hoge kwaliteit data**
De datakwaliteit bepaald in hoeverre de resultaten kloppen met de werkelijkheid
- ✓ **Innovatief dashboard**
Met een innovatief dashboard krijgt de inkoopafdeling meer inzicht op basis van het dashboard.



Doel van de workshop

5



5 minuten

Dashboard demo

6



5-10 minuten



20 minuten

Huidige situatie

8



Hoe analyseren we

Welke analyses maken we

Welke data gebruiken we nu

Wat zijn de knelpunten

30 minuten

KPI's definiëren

SMART

1

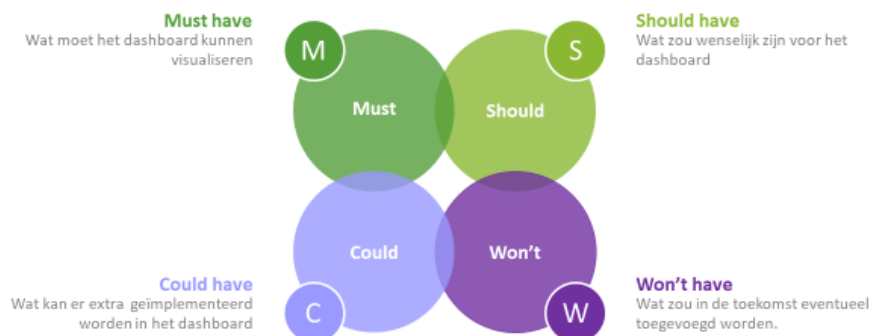


20 minuten

KPI's prioriteren

MoSCoW methode

10



30 minuten

Requirements

11

Als <gebruiker/rol> wil ik <functie/actie>, zodat ik <doel>.

Voorbeelden:

- Als een nieuws-redacteur wil ik de artikelen reviewen voordat ze gepubliceerd worden, zodat ik kan verzekeren dat de artikelen geen spelfouten meer bevatten.
- Als student, wil ik online toegang hebben tot mijn rooster, zodat ik weet welke lessen ik op welk moment, op welke locatie heb.
- Als inkoper wil ik kunnen aflezen welke leveranciers voorkeurleveranciers zijn, zodat ik de meest geschikte leverancier kan benaderen voor mijn doeleinde.
- Als inkoper wil ik factuurdata met inkoopdata combineren, zodat ik een complete beeld krijg over de inkoopstatus.

30 minuten

Requirement Prioriteren

12





Roadmap



Zijn er nog vragen?

10.4 Bijlage 4 Datakwaliteit SQL scripts

```
15 TRUNCATE TABLE [dbo].[datacontrole]
16 TRUNCATE TABLE [dbo].[dataerror]
17
18
19 DECLARE @KolCur varchar(100)
20 ,@SQLStatement varchar(max)
21 ,@CodeStatement varchar(max)
22 ,@Kolom varchar(100)
23 ,@Tabel varchar(50)
24 ,@Where varchar(max)
25 ,@WhereState varchar(max)
26 ,@Code varchar(max)
27 ,@Omschrijving varchar(max)
28
29
30 DECLARE Kolom_Cursor CURSOR LOCAL FOR
31
32 SELECT DISTINCT table_name + '/' + column_name as KolCur
33 FROM information_schema.columns
34 WHERE 1 = 1
35 AND table_name in ('dim_leverancier', 'dim_leverancierkenmerken')
36 AND COLUMN_NAME in ('administratie', 'status', 'bestelmedium',
37 'emailadres', 'betalingsconditie', 'betalingsconditie_crediteur',
38 'betalingstermijn_crediteur_dagen', 'kernleverancier', 'leveringsconditie',
39 'naam', 'wka_leverancier', 'admin_code', 'soort', 'kenmerk', 'omschrijving',
40 'waarde', 'datum_geldig_van', 'datum_geldig_tem'
41 )
42 ORDER BY KolCur
43
44 OPEN Kolom_Cursor
45
46 FETCH NEXT FROM Kolom_Cursor INTO @KolCur
47 WHILE (@@FETCH_STATUS = 0)
48 BEGIN
49
50 SET @Tabel = SUBSTRING(@KolCur, 0, CHARINDEX('/', @KolCur, 0))
51 SET @Kolom = SUBSTRING(@KolCur, CHARINDEX('/', @KolCur, 0) + 2, LEN(@KolCur) - CHARINDEX('/', @KolCur, 0))
52 SET @Code = CASE WHEN @Tabel = 'dim_leverancier' THEN '[administratie]'
53 WHEN @Tabel = 'dim_leverancierkenmerken' THEN '[admin_code]'
54 ELSE 'Failed'
55 END
56 SET @Omschrijving = CASE WHEN @Tabel = 'dim_leverancier' THEN '[leverancier_skey]'
57 WHEN @Tabel = 'dim_leverancierkenmerken' THEN '[leverancierkenmerken_skey]'
58 ELSE 'Niet beschikbaar'
59 END
```




```
60 SET @where = CASE --WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'Administratie' THEN ' WHERE dim_leverancier.Administratie = 'B939'' --TEST query
61 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'status' THEN ' WHERE [Status] NOT IN ('Geblokkeerd', 'Vrij voor gebruiker')'
62 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'emailadres' THEN ' WHERE bestelmedium LIKE '308' AND
63 emailadres NOT LIKE '%[A-Za-z]%' AND
64 status = 'Vrij voor gebruiker' AND
65 naam NOT LIKE '%Croonwolder&dros%'
66 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'administratie' THEN ' WHERE [administratie] not like 'B9%' and [administratie] not like 'B1%'
67 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'bestelmedium' THEN ' WHERE Bestelmedium NOT IN
68 (
69 SELECT DISTINCT Bestelmedium
70 FROM dim_leverancier
71 WHERE [Administratie] LIKE 'B110'
72 )'
73 --Wanneer ik twee queries op één column plaats, vervangt de laatste query de eerste
74 --betalingsconditie mist de dubbele codes die niet gelijk zijn aan de B110
75 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'betalingsconditie' THEN ' WHERE betalingsconditie NOT IN
76 (
77 SELECT DISTINCT betalingsconditie
78 FROM dim_leverancier
79 WHERE [Administratie] LIKE 'B110'
80 )'
81 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'betalingsconditie_crediteur' THEN ' WHERE betalingsconditie_crediteur NOT IN
82 (
83 SELECT DISTINCT betalingsconditie_crediteur
84 FROM dim_leverancier
85 WHERE [Administratie] LIKE 'B110'
86 )'
87 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'betalingstermijn_crediteur_dagen' THEN ' WHERE [betalingstermijn_crediteur_dagen] = 0 AND
88 betalingsconditie_crediteur <> 'B /
89 Betaling geblokkeerd'
90 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'kernleverancier' THEN ' WHERE [Kernleverancier] NOT LIKE 'Ja' AND [Kernleverancier] NOT LIKE 'Nee'
91 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'leveringsconditie' THEN ' WHERE [leveringsconditie] = '0'
92 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'naam' THEN ' WHERE naam not like '%[A-Za-z]%'
93 WHEN @Tabel = 'dim_leverancier' AND @Kolom = 'wka_leverancier' THEN ' WHERE [wka_leverancier] NOT LIKE 'J' AND [wka_leverancier] NOT LIKE 'N'
94 --Leverancierkenmerken test enkel op Administratie B110.
95 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'admin_code' THEN ' WHERE admin_code NOT LIKE 'B9%' and admin_code not like 'B1%'
96 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'soort' THEN ' WHERE admin_code = 'B110' AND soort NOT LIKE '%[A-Za-z]%'
97 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'kenmerk' THEN ' WHERE admin_code = 'B110' AND kenmerk NOT LIKE '%[A-Za-z]%'
98 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'omschrijving' THEN ' WHERE admin_code = 'B110' AND omschrijving NOT LIKE '%[A-Za-z]%'
99 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'waarde' THEN ' WHERE admin_code = 'B110' AND waarde NOT LIKE '%[A-Za-z]%'
100 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'datum_geldig_van' THEN ' WHERE admin_code LIKE 'B110'
101 AND datum_geldig_van BETWEEN DateAdd(YEAR, -20, GetDate())
102 AND DateAdd(YEAR, -5, GetDate())
103 WHEN @Tabel = 'dim_leverancierkenmerken' AND @Kolom = 'datum_geldig_tem' THEN ' WHERE admin_code = 'B110'
104 AND datum_geldig_tem BETWEEN DateAdd(YEAR, -100, GetDate())
105 AND DateAdd(YEAR, -5, GetDate())
106 OR admin_code = 'B110' AND datum_geldig_tem > DateAdd(YEAR, +5, GetDate())
107 ELSE ''
108 END
109 SET @whereState = ''
110 SET @whereState = CONVERT(VARCHAR(1000), @where)
111 SET @whereState = REPLACE(@where, ' ', '')
112 SET @whereState = replace(@whereState, '\t', '');
113
114 -- INSERT where statements INTO _datacontrole
115 SET @SQLstatement = ''
116 SET @SQLstatement =
117 '
118 INSERT INTO [dbo].[_datacontrole]
119 ([Tabel]
120 , [Kolom]
121 , [Conditie]
122 , [Aantal])
123 SELECT '' + @Tabel + '' as Tabel, '' + @Kolom + '' as Kolom, ''Where'' as Conditie, Count(' + @Kolom + ') as cnt_ + @Tabel + ' + @Kolom + '
124 FROM '' + @Tabel + '' +
125 @where
126
127 EXEC (@SQLstatement) -- Execute Dynamic SQL statement
128 --select (@SQLstatement) -- Execute Dynamic SQL statement
129
130
131
132 --INSERT rowcount of column into _datacontrole
133 SET @SQLstatement = ''
134 SET @SQLstatement =
135 '
136 SELECT '' + @Tabel + '' as Tabel, '' + @Kolom + '' as Kolom, ''Totaal'' as Conditie, Count(' + @Kolom + ') as cnt_ + @Tabel + ' + @Kolom + '
137 FROM '' + @Tabel
138
139 EXEC (@SQLstatement) -- Execute Dynamic SQL statement
140 --select (@SQLstatement) -- Execute Dynamic SQL statement
141
142
```



```
143 -----
144 --DATAERROR TABEL
145 -----
146
147 SET @CodeStatement = ''
148 SET @CodeStatement =
149 '
150 INSERT INTO dbo._dataerror (
151     [Tabel]
152     , [Kolom]
153     , [Code]
154     , [Conditie]
155     , [Omschrijving]
156     , [Waarde]
157 )
158 SELECT ''' + @Tabel + ''' as Tabel,
159        ''' + @Kolom + ''' as Kolom,
160        ' + @Code + ' as Code,
161        ''' + TRIM(@WhereState) + ''' as Conditie,
162        ' + @Omschrijving + ' as Omschrijving,
163        ' + @Kolom + ' as Waarde
164 FROM ' + @Tabel
165        + @Where
166
167
168 EXEC (@CodeStatement)
169 --SELECT @CodeStatement
170
171
172
173 FETCH NEXT FROM Kolom_Cursor INTO @KolCur
174 END
```



10.5 Bijlage 5 meetresultaten Excel en Power BI

	tab	id	gew	data rules	bevestiging	auto increment	record	aantal	aantal	aantal foutieve waarden	aantal dubbele waarden	aantal distinct	aantal gouvernante	testwijze
d. leverancier	leverancier_skey	3			Auto increment hoeft niet geëst te worden Check of administraties leeg gelaten worden of niet voorkomen in "leverancier", maar wel in andere tabellen	True	113906	0	0	0	0	113906	113906	Auto increment Een measure gemaakt "Aantal B800 administraties" die filtert op B800 administraties.
d. leverancier	Administratie	B800 series worden geen leveranciers ingelazen. 2 B800 series is fout.			B110 is ljdend, verander de verschillende omschrijvingen naar B110 administratie. Check vanaf 2017.	113906	0	0	0	0	6393	16	14	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Bestelmedium	De omschrijving moet gelijk zijn aan B110 omschrijving			Zit een dubbele waarde in en bevat een 0 waarde	113906	0	0	0	0	704	14	13	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Bestellingconditie	2 0 is niet fout, leeg en dubbele waarden wel mag niet leeg gelaten worden, dubbel moeten 2 hetzelfde zijn. 0 is niet fout.			Zit een dubbele waarde in en bevat waarden met de filter op "V" toont betalingsconditie_crediteur 4 verschillende waarden.	113906	2865	0	17	13	13	13	13	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Bestellingtermijn_crediteur_dagen	Verschiende 0 waarden zijn incorrect, leeg is fout 2 en mag geen dubbele waarden bevatten.			Met de filter op "V" toont betalingsconditie_crediteur 4 verschillende waarden.	113906	0	10007	0	3	0	3	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Emailadres	Als de betalingsmedium via e-mail is, moet er een 2 emailadres zijn. Andere gevallen is leeg niet fout.			Alleen belangrijk wanneer bestelmedium email is	113906	87	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Kennelijkverander	3 mag ja/nee zijn 0 waarden is fout, in dat geval wordt het veld leeg			100% correct	113906	37	0	0	2	2	2	2	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Leveringsconditie	1 gelaten			bevat veel 0 waarden	113906	0	5841	0	12	12	12	12	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Naam	2 Naam mag niet leeg gelaten worden			Bevat lege en foutieve waarden	113906	2835	5	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Status	2 is verplicht en mag niet leeg gelaten worden			Correct	113906	0	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Mkt leverancier	1 mag niet leeg zijn			Correct	113906	0	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Leverancierskey	1 Email B110 en B800 series bevatten			Auto increment, hoeft niet geëst te worden	True	775318	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Admin_skey	2 Wordt automatisch ingevoerd, het op B110 is de den is correct, later bepalen hoe de B929 series gelijk is. True			B800 series is wakkend	775318	0	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Leverancierskey	2 Automatisch			Auto increment	True	775318	-	-	-	-	-	-	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Labelcode	1 0 is fout			Waarom staan er 0 waarden in, de soort, kenmerk en waarde velden zijn leeg. Wordt niet gebruikt, maar kan veel redundante data zijn.	775318	0	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Soort	2 Blanks is fout			Bevat nul waarden, wanneer soort p van een raamcontract. Enkel de B929 bevat blanks. Wellicht gaat er iets fout bij het laden.	775318	0	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Kenmerk	1 Blanks is fout, verder alles goed			Bevat nul waarden	775318	1758	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Omschrijving	1 Blanks is fout, verder alles goed			Bevat nul waarden, is dit belangrijk	775318	12	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Waarde	1 Blank is fout			Besprek met de BT manager of blanks en geen hetzelfde is of dat blanks als fout gezien moet worden	775318	10576	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Datum geldig van	2 1901 is fout			Bevat outliers naar 2050 en 1901, is dit belangrijk?	775318	4129	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.
d. leverancier	Datum geldig tot en met	2 1901 is fout. Maximaal 5 jaar in de toekomst, alles 2 daarboven is fout.			Bevat outliers naar 2050 en 1901, is dit belangrijk.	775318	386	0	0	0	0	0	0	De foutieve waarden zijn opgeteld door de volgende filters te toe te passen: "Incasso" 10 en "9/ Betaling 21 dagen". Wanneer de filter op "V" staat, mag betalingsconditie_crediteur maar één 1 distinct bevatten.

10.6 Bijlage 6 meetresultaten SQL Server

```

1  /***** Script for SelectTopNRows command from SSMS *****/
2  SELECT [Tabel]
3         , [Kolom]
4         , [Conditie]
5         , [Aantal]
6  FROM [dbo].[_datacontrole]
7

```

100 %

Results Messages

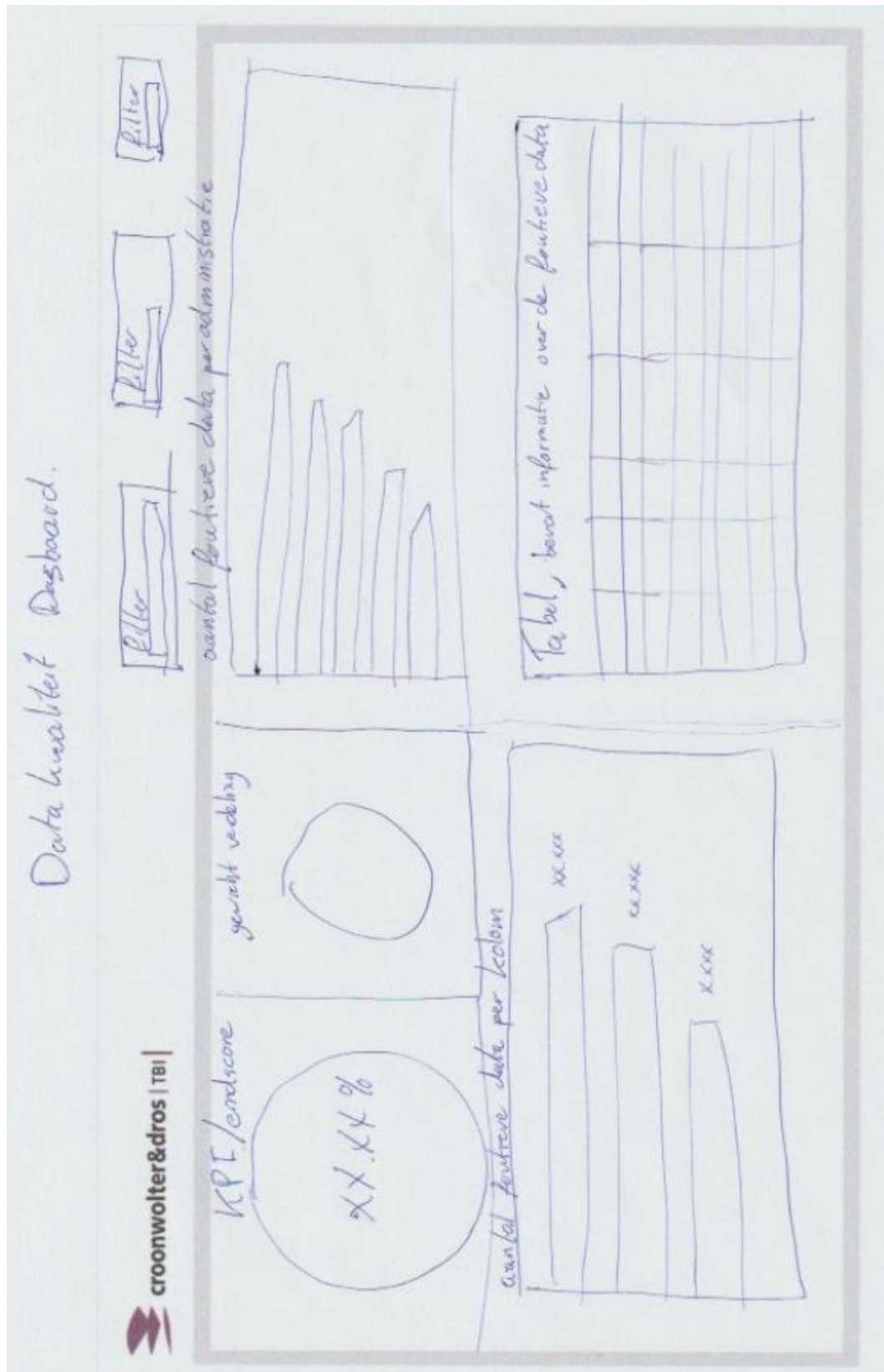
	Tabel	Kolom	Conditie	Aantal
1	dim_leverancier	administratie	Where	0
2	dim_leverancier	bestelmedium	Where	6309
3	dim_leverancier	betalingsconditie	Where	704
4	dim_leverancier	betalingsconditie_crediteur	Where	2865
5	dim_leverancier	betalingstemijn_crediteur_dagen	Where	10007
6	dim_leverancier	emailadres	Where	87
7	dim_leverancier	kemleverancier	Where	37
8	dim_leverancier	leveringsconditie	Where	5841
9	dim_leverancier	naam	Where	2840
10	dim_leverancier	status	Where	0
11	dim_leverancier	wka_leverancier	Where	0
12	dim_leverancierkenmerken	admin_code	Where	0
13	dim_leverancierkenmerken	datum_geldig_tem	Where	386
14	dim_leverancierkenmerken	datum_geldig_van	Where	4129
15	dim_leverancierkenmerken	kenmerk	Where	1758
16	dim_leverancierkenmerken	omschrijving	Where	12
17	dim_leverancierkenmerken	soort	Where	0
18	dim_leverancierkenmerken	waarde	Where	10576

✓ Query executed successfully.

10.7 Bijlage 7 DAX Statements formule gestandaardiseerd datakwaliteit cijfer

Kolom/Measure naam	Type	DAX Statement
Percentage	Kolom	Percentage = SWITCH(TRUE(); _datacontrole_administratie[tabel] = "dim_leverancierkenmerken"; DIVIDE(_datacontrole_administratie[Aantal];[rowcount leverancierkenmerken]); _datacontrole_administratie[tabel] = "dim_leverancier"; DIVIDE(_datacontrole_administratie[Aantal];[rowcount leverancierkenmerken]))
Gemiddelde	Kolom	gemiddelde = SWITCH(TRUE(); _datacontrole_administratie[tabel] = "dim_leverancier"; DIVIDE(RELATED(d_leverancier_gewicht[Gewicht]);SUM(d_leverancier_g ewicht[Gewicht])); _datacontrole_administratie[tabel] = "dim_leverancierkenmerken"; DIVIDE(RELATED(d_leverancier_gewicht[Gewicht]);SUM(d_leverancier_g ewicht[Gewicht])))
Percentage * gewicht	Kolom	percentage x gewicht = _datacontrole_administratie[percentage] * RELATED(d_leverancier_gewicht[Gewicht])
Datakwalitei t cijfer	Measure	datacontrole kwaliteitcijfer per administratie = 1 - SUM([percentage x gewicht]) / CALCULATE(SUM(d_leverancier_gewicht[Gewicht]);ALL(d_leverancier_ge wicht[Gewicht]))
Rowcount	Measure	rowcount leverancierkenmerken = COUNTROWS(d_leverancierkenmerken) rowcount leverancier = COUNTROWS(d_leverancier)

10.8 Bijlage 8 Blauwdruk schets Datakwaliteit dashboard

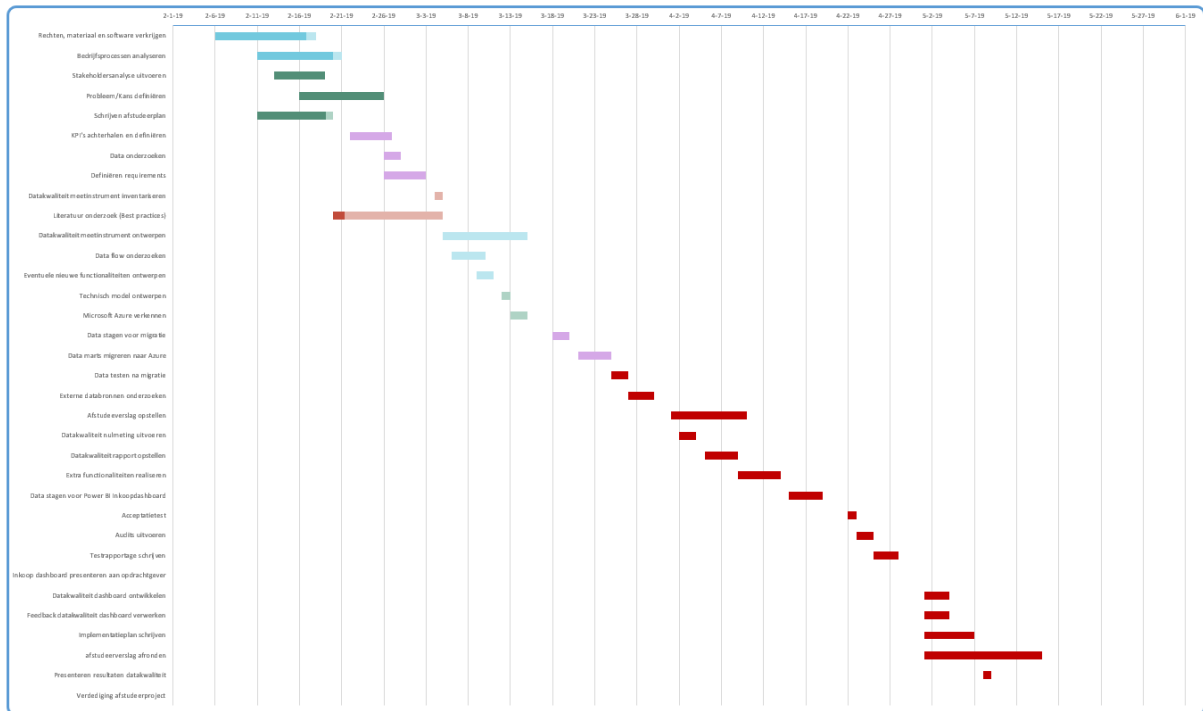


Figuur 27 DQ dashboard schets

10.9 Bijlage 9 planning

ITERATIES	ACTIVITEITEN
VOORONDERZOEK 12-02-2019 - 26-02-2019	Rechten, materiaal en software verkrijgen. Bedrijfsprocessen analyseren Stakeholderanalyse uitvoeren Probleem/kans definiëren Schrijven afstudeerplan
DEFINIËREN 26-02-2019 – 12-03-2019	KPI's achterhalen en definiëren Definiëren requirements Data onderzoeken Datakwaliteit meetinstrument inventariseren Literatuur onderzoek (best practices)
ITERATIE 1 13-03-2019 - 25-03-2019	Data management meetinstrument ontwerpen Data flow onderzoeken Data management methodieken onderzoeken Microsoft Azure verkennen
ITERATIE 2 26-03-2019 – 03-04-2019	Data management meetinstrument testen Microsoft Azure datakwaliteit services onderzoeken Eventuele nieuwe functionaliteiten onderzoeken Externe databronnen onderzoeken
ITERATIE 3 04-04-2019 – 19-04-2019	Afstudeerverslag opstellen Data management nulmeting uitvoeren Data management rapport opstellen Extra functionaliteiten realiseren Data stagen voor Power BI inkoopdashboard
ITERATIE 4 22-04-2019 – 30-04-2019	Acceptatietest Audits uitvoeren Testrapportage schrijven Inkoop dashboard presenteren aan opdrachtgever
NAZORG 01-05-2019 – 15-05-2019	Feedback inkoop dashboard verwerken Implementatieplan schrijven Afstudeerverslag afronden Presenteren resultaten datakwaliteit Verdediging afstudeerproject Datakwaliteit dashboard ontwikkelen

Iteratie	Task Name	Start Date	End Date	Duration (Days)	Days Complete	Days Remaining	Percent Complete
1	Rechten, materiaal en software verkrijgen	06-02-2019	18-02-2019	12	10,80	1,20	90%
1	Bedrijfsprocessen analyseren	11-02-2019	21-02-2019	10	9,00	1,00	90%
1	Stakeholdersanalyse uitvoeren	13-02-2019	19-02-2019	6	6,00	0,00	100%
1	Probleem/Kans definiëren	16-02-2019	26-02-2019	10	10,00	0,00	100%
1	Schrijven afstudeerplan	11-02-2019	20-02-2019	9	8,10	0,90	90%
2	KPI's achterhalen en definiëren	22-02-2019	27-02-2019	5	0,00	5,00	0%
2	Data onderzoeken	26-02-2019	28-02-2019	2	0,00	2,00	0%
2	Definiëren requirements	26-02-2019	03-03-2019	5	0,00	5,00	0%
2	Datakwaliteit meetinstrument inventariseren	04-03-2019	05-03-2019	1	0,00	1,00	0%
2	Literatuur onderzoek (Best practices)	20-02-2019	05-03-2019	13	1,30	11,70	10%
3	Datakwaliteit meetinstrument ontwerpen	05-03-2019	15-03-2019	10	0,00	10,00	0%
3	Data flow onderzoeken	06-03-2019	10-03-2019	4	0,00	4,00	0%
3	Eventuele nieuwe functionaliteiten ontwerpen	09-03-2019	11-03-2019	2	0,00	2,00	0%
3	Technisch model ontwerpen	12-03-2019	13-03-2019	1	0,00	1,00	0%
3	Microsoft Azure verkennen	13-03-2019	15-03-2019	2	0,00	2,00	0%
4	Data stagen voor migratie	18-03-2019	20-03-2019	2	0,00	2,00	0%
4	Data marts migreren naar Azure	21-03-2019	25-03-2019	4	0,00	4,00	0%
4	Data testen na migratie	25-03-2019	27-03-2019	2	0,00	2,00	
4	Externe databronnen onderzoeken	27-03-2019	30-03-2019	3	0,00	3,00	
5	Afstudeeverslag opstellen	01-04-2019	10-04-2019	9	0,00	9,00	
5	Datakwaliteit nulmeting uitvoeren	02-04-2019	04-04-2019	2	0,00	2,00	
5	Datakwaliteit rapport opstellen	05-04-2019	09-04-2019	4	0,00	4,00	
5	Extra functionaliteiten realiseren	09-04-2019	14-04-2019	5	0,00	5,00	
5	Data stagen voor Power BI Inkoopdashboard	15-04-2019	19-04-2019	4	0,00	4,00	
6	Acceptatietest	22-04-2019	23-04-2019	1	0,00	1,00	
6	Audits uitvoeren	23-04-2019	25-04-2019	2	0,00	2,00	
6	Testrapportage schrijven	25-04-2019	28-04-2019	3	0,00	3,00	
6	Inkoop dashboard presenteren aan opdrachtgever	30-04-2019	30-04-2019	0	0,00	0,00	
7	Datakwaliteit dashboard ontwikkelen	01-05-2019	04-05-2019	3	0,00	3,00	
7	Feedback datakwaliteit dashboard verwerken	01-05-2019	04-05-2019	3	0,00	3,00	
7	Implementatieplan schrijven	01-05-2019	07-05-2019	6	0,00	6,00	
7	afstudeerverslag afronden	01-05-2019	15-05-2019	14	0,00	14,00	
7	Presenteren resultaten datakwaliteit	08-05-2019	09-05-2019	1	0,00	1,00	
7	Verdediging afstudeerproject	31-05-2019	31-05-2019	0	0,00	0,00	



10.10 Bijlage 10 management presentatie



Data Driven inkoop

Inhoud

2



- ✓ **Introductie**
Wat is de opdracht inclusief de scope
- ✓ **Huidige Status**
Wat is gedaan en wat moet er nog gebeuren
- ✓ **Afsluiting**
Tijd voor vragen en opmerkingen



Doel van de workshop

3

Wie zijn we

- Sike van Dongen
- Dennis Loos



Wat doen we

- Datakwaliteit meten inkoopafdeling
- User Experience optimaliseren
- Inkoop dashboard ontwikkelen

Wat is ons doel

Inkoopafdeling voorzien van een gebruiksvriendelijk dashboard waarmee beslissingen kunnen gemaakt worden op basis van betrouwbare data.

Route

4

Huidige status

- Datakwaliteit leverancier
- Dashboard leverancier
- Scrum proces
 - Sprints
 - Reviews
 - Product backlog

To-Do

- Inkoop dashboard door ontwikkelen
- Datakwaliteit meten
- Externe/Open data onderzoek
- Datakwaliteit dashboard
- Inkoop dashboard Testen

Vaststellen

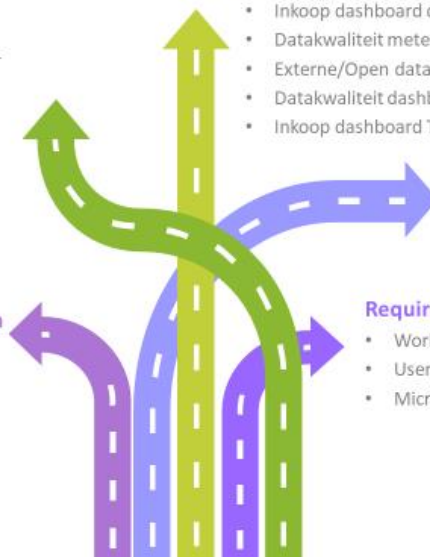
- Definities
- Targets
- Relevante databronnen

Opdracht vormgeven

- Probleemanalyse
- Stakeholderanalyse
- Risicoanalyse
- Project organisatie
- Planning

Requirements analyse

- Workshop Inkoopers
- User Stories
- Microsoft Azure Dev Ops opzetten



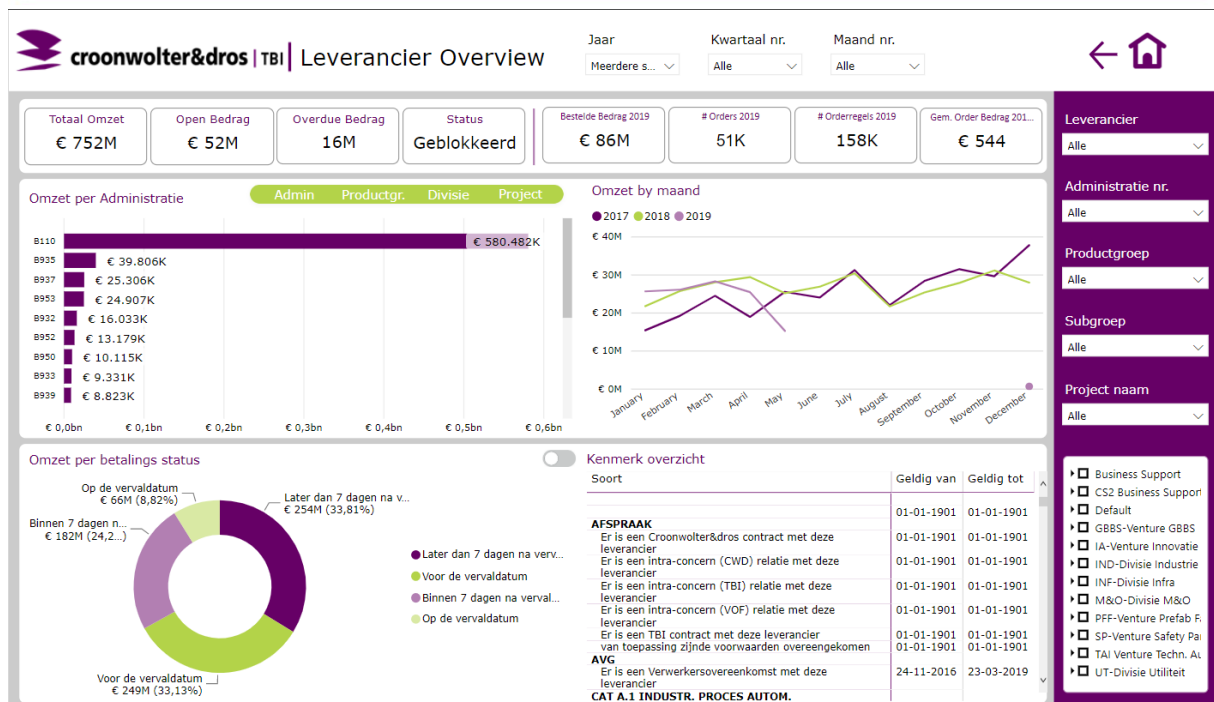
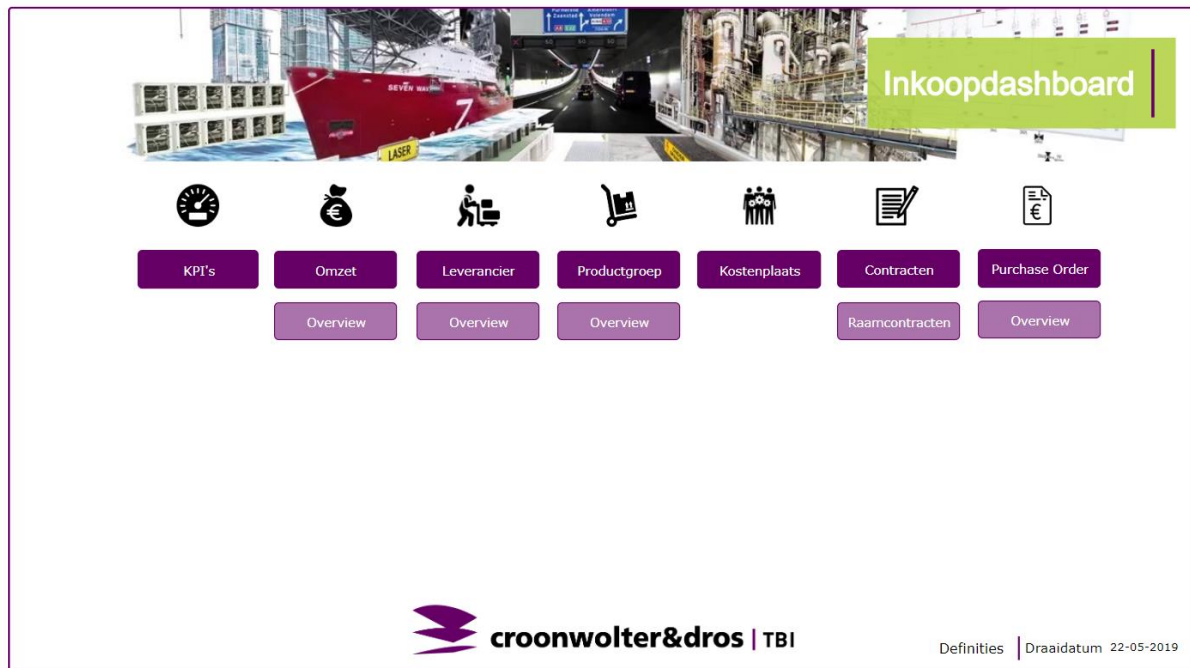
Key Performance Indicators

- Aantal bestelorders dat via raamcontracten is ingekocht
- Aantal bestellingen dat met een bestelorder is gemaakt
- Aantal openstaande facturen per leverancier
- De facturomzet van een Administratie
- De facturomzet van inkopen via VKL vs. niet-VKL
- Aantal leveranciers die open zijn voor gebruik vs. totaal aantal leveranciers



Zijn er nog vragen?

10.10.1 Demo Croonwolder&dros inkoopdashboard



10.11 Bijlage 11 meeting factuurregels zonder kostenplaats

Meeting 18-04-2019

Aanwezig: Paul Batenburg, Sike v Dongen en Dennis Loos

Facturen die niet te herleiden zijn aan een project nummer en/of kostenplaats

De factuurnummers en boekingsnummers komen in veel gevallen overeen in Power BI. Er zijn veel velden die een factuurnummer bevatten, maar geen boekingsnummer. De velden zonder boekingsnummer zijn niet terug te vinden in Acto.

- 0 cent facturen beginnen altijd met 98
- UVA-Batch = Uren Verantwoording Batch, UVA is voor iedereen die uren heeft gemaakt, interne en externe werknemers. UVA batches staan in relatie tot een project
- Als er geen project is gedefinieerd voor een factuur, moet er een kostenplaats ingevuld zijn
- Facturen met een mis match kun je op kostenplaats filteren, ook niet-order facturen

In het orderregelniveau zit het project nummer, zie onderstaande afbeelding van Acto. Het projectnummer is hieruit te herleiden.

Acto Hoofdmenu Inleen Wijzigen en verzenden bestelorders (LPM201)

Zoeken in 'Wijzigen en verzenden bestelorders (LPM201 / PBATEN / PBATEN_201 / aixprod1 / P / B110 / 19.10)'; alternatief=INLEEN+

Bestand Start Bewerken Zoeken Extra Speciaal Help

Alternatief Actualiseren Zoeken Kopieren Exporteren CSV_Export Sorteren ... Herschikken Kolombreedtes Mark_Omkieren Mark_Alles Mark_Geen Analyseren Bovenaan PageUp MoveUp MoveDo

1 van 8 Tekst zoeken

Batchnr.	UVA	Leverancier	Bestelorde	Bestelp	Bestelreg	Projectnummer	Vestig	Naam	Aanvrager	Bruto	Besteldat	Status
								vestiging		bedrag		
21912001	TECLINE	23294821	01	01	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		1.360,00	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	02	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		214,20	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	03	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		238,00	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	04	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		1.360,00	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	05	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		142,80	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	06	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		238,00	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	07	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		1.360,00	27/03/2019	3/Geprint
21912001	TECLINE	23294821	01	08	21710100	175	Amsterdam - Hettenheuvelweg 50	Kruik J.P.		1.360,00	27/03/2019	3/Geprint

Sommige bestelorderregels zijn niet te herleiden naar een projectnummer of vestiging, terwijl andere regels in dezelfde bestelorder dit wel het geval is.

Zoeken in 'Wijzigen en verzenden bestelorders (LPM201 / PBATEN / PBATEN_201 / aixprod1 / P / B110 / 19.10)'; alternatief=INLEEN+

Bestand Start Bewerken Zoeken Extra Speciaal Help

Alternatief Actualiseren Zoeken Kopieren Exporteren CSV_Export Sorteren ... Herschikken Kolombreedtes Mark_Omkieren Mark_Alles Mark_Geen Analyseren Bovenaan PageUp MoveUp MoveDo

3 van 6 Tekst zoeken

Batchnr.	UVA	Leverancier	Bestelorde	Bestelp	Bestelreg	Projectnummer	Vestig	Naam	Aanvrager	Bruto	Besteldat	Status
								vestiging		bedrag		
81912001	PMC ZCP	23294443	01	05	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.		66,96	27/03/2019	3/Geprint
81912001	PMC ZCP	23294443	01	04	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.		144,77	27/03/2019	3/Geprint
81912001	PMC ZCP	23294443	01	06	00000000	000		Zeeuw de M.S.		247,26	27/03/2019	3/Geprint
81912001	PMC ZCP	23294443	01	02	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.		1.287,80	27/03/2019	3/Geprint
81912001	PMC ZCP	23294443	01	01	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.		2.833,16	27/03/2019	3/Geprint
81912001	PMC ZCP	23294443	01	03	40001807	400	IND-Divisie Industrie	Zeeuw de M.S.		128,78	27/03/2019	3/Geprint



Met het zoeken met de Batchnummer om het project nummer en/of de kostenplaats te achterhalen toont Acto het volgende scherm: Ook hier ontbreekt het project nummer en staat er geen kostenplaats.

Acto

Hoofdmenu

Urenverantwoording

Inbrengen/corrigeren urenverantwoordingen (PRM568)

Inbrengen/corrigeren urenverantwoordingen (PRM568 / PBATEN / PBATEN_201 / aikprod1 / P / B110 / 19.10)

Bestand

Start

Bewerken

Zoeken

Extra

Speciaal

Vensters

Help

PgUp

PgDn

Toevoegen

Uitbreiden

Wijzigen

Verwijder

Rapporteren

Archiveren

Samenvoegen

Zoeken

Vorige

Eerstvolgende

Zactualiseren

Zoekbox...

Zvorige

Zeerstvolgende

Taakmonitor

FAQ

Opdracht

Batchnr

81912001

(Hoofdwjkr 12 / 2019)

Jaarweek

2019 / 12

100,00 %

Pers nummer

734130 Backhuis, E.J.

Dir: 32,00

Stand: 40,00

Overuren: 1,00

Soort wkr

70 Int'proj.leidmanag.

00004170

Ind: 0,00

Rest: 8,00

Reisuren: 0,00

Regel nummer	Project nummer	Project naam		Kosten plaats	Activiteit nummer	Activiteit naam	Uren soort	Vestig numm	Werk uren	Ploegen toeslag	Reis uren	Reis kilometer	code	aantal	Over uren	code	aantal	code	aan code	code	Indir'ect1 dag	saldo-code	Indir'ect2 dag	saldo	code	Onk'osten bedrag	code	bedrag
10	400.01807	Matrix project control	P		1004	BESIX BESIX	41	108	0,00		0,00	0	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	37	66,96	0	0,00
11	400.01807	Matrix project control	P		1012	TENNET TEN TEN	41	108	0,00		0,00	0	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	37	144,77	0	0,00
12	000.00000		-					108	0,00		0,00	0	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	136	247,26	0	0,00
13	400.01807	Matrix project control	P		1004	BESIX BESIX	41	108	10,00		0,00	0	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	0	0,00	0	0,00
14	400.01807	Matrix project control	P		1012	TENNET TEN TEN	41	108	22,00		0,00	0	0	0,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	0	0,00	0	0,00
15	400.01807	Matrix project control	P		1012	TENNET TEN TEN	41	108	0,00		0,00	0	1	1,00	0	0,00	0	0,00	0	0,00	0	0,00	0,00	0,00	0	0,00	0	0,00

Om de kostenplaats te kunnen herleiden moet er gezocht worden via het boekstuknummer in het historische grootboek:

Zoeken in Historisch grootboek (globaal) (FIN052 / PBATEN / PBATEN_201 / aikprod1 / P / B110 / 19.10), alternatief=SNELLER																
Bestand Start Bewerken Zoeken Extra Speciaal Help																
Rapporteren Alternatief Actualiseren Zoeken Kopieren Exporteren CSV_Export Sorteren ... Herschikken Kolonbreedtes Mark_Omkeren Mark_Alles Mark_Geen Analyseren Bovenaan PageUp MoveUp MoveDown PageDown Onderaan																
12 van 12 Tekst zoeken Bron actuele gegevens																
Grootboek- nummer	Grootboek- omschr	Kostenpl- nummer	Kostenpl- omschr	Periode- nummer	81912001						'734130'					
					Dagboek	Boekstuk- nummer	Volgnr	Aantal	Bedrag	D-C	Boekdata	Omschrijving	Project	Proj.de	B.nr	Batchr
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042660	0,00	247,26	C	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042678	0,00	66,96	C	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042679	10,00	1.287,80	C	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042680	0,00	144,77	C	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042681	22,00	2.833,16	D	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00015640 Nog te ontf. inleners	00000000	Default	201903	0035	81912001	46042682	1,00	128,78	C	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982
✓	00032040 Inleen eng.montage-PL-PM	L	00004000 IND-Divisie Industrie	201903	0035	81912001	46042693	0,00	66,96	D	27/03/2019	Backhuis, /734130/12/2019	40001807	00000	81912001	6982
✓	00032040 Inleen eng.montage-PL-PM	L	00004000 IND-Divisie Industrie	201903	0035	81912001	46042694	10,00	1.287,80	D	27/03/2019	Backhuis, /734130/12/2019	40001807	00000	81912001	6982
✓	00032040 Inleen eng.montage-PL-PM	L	00004000 IND-Divisie Industrie	201903	0035	81912001	46042695	0,00	144,77	D	27/03/2019	Backhuis, /734130/12/2019	40001807	00000	81912001	6982
✓	00032040 Inleen eng.montage-PL-PM	L	00004000 IND-Divisie Industrie	201903	0035	81912001	46042696	22,00	2.833,16	D	27/03/2019	Backhuis, /734130/12/2019	40001807	00000	81912001	6982
✓	00032040 Inleen eng.montage-PL-PM	L	00004000 IND-Divisie Industrie	201903	0035	81912001	46042697	1,00	128,78	D	27/03/2019	Backhuis, /734130/12/2019	40001807	00000	81912001	6982
✓	00040900 Inleen indirect personeel		00004170 IND-BUOP-Realisatie	201903	0035	81912001	46042734	0,00	247,26	D	27/03/2019	Backhuis, /734130/12/2019	00000000	00000	81912001	6982

Op het dashboard dat in ontwikkeling is voor de Inkoop afdeling geeft factuurnummer/boekingsnummer geen kostenplaats en geen projectnummer aan. In Acto is dit wel genoteerd. Zie onderstaande afbeelding

Inkoopfacturen (INK150 / PBATEN / PBATEN_201 / aikprod1 / P / B110 / 19.10)																						
Bestand		Start	Bewerken	Zoeken	Extra	Speciaal	Vensters	Help														
PgUp		PgDn	Ga naar	Toevoegen	Uitbreiden	Wijzigen	Kolom	Verwijder	Rapporteren	Archiveren	Samenvoegen	Zoeken	Vorige	Eerstvolgende	Zactualiseren	Zoekbox...	Zvorige	Zeerstvolgende	Taakmonitor			
Crediteur 464443 Interparking Nederland B.V.									Status Verwerkt													
Boekstuknummer 01010744									Bestelorder(s) 00000000 Besteld													
Factuurnr lev. 0002200299									Factuurdatum 01/01/2019 Ontvangen													
Factuurbedrag 503,41 EUR									Geboekt bedrag 416,04 Gefactureerd													
Btw-bedrag 87,37 EUR									Nog te boeken 0,00 Oopenstaand													
1 Algemeen									Opdracht													
2 Regels																						
Markering	Rg	Bestel	Best	Beste	Boeking	Grootboek	Artikel	Aantal	Bedrag	Aantal	Aantal	Aantal	Project	Pe	Co	Activiteit	Kosten	Kosten	Budget	Bw		
	nr	order	post	regel	soort	nummer	nummer	factuur	factuur	besteld	ontvangen	gefactureerd	nummer	nu	lin	v	deel	soort	plaats	bedrag	code	
☐	1					9302	32100	PR K Z O INTERPARKI	0,00	416,04			257.10016				00000	5300	00002570	03		
☐																						

10.12 Bijlage 12 Assessment data mart crediteuren

ACTO meeting Inkoop data

Document eigenaar: Dennis Loos

Gecontroleerd door: Susan Verbeij

Datum 12-02-2019

Locatie Croonwolter&dros Amersfoort

Deelnemers

NAAM

Susan Verbeij

Dennis Loos

FUNCTIE

Inkoper - CW&D Staff-inkoop-Support

Projectleider - Stagiair

Op basis het assessment van de data marts uit ACTO zijn er een aantal vragen en opmerkingen opgesteld. Onderstaande tabel omschrijft de vragen of opmerkingen die zijn opgevallen tijdens het assessment van de Acto data marts:

Opmerkingen- en vragentabel

Tabelnaam	Opmerkingen
D_project	- De tabel d_project bevat kolommen met super en mega projecten. Wat is de definitie hiervan en wanneer wordt iets als een super of mega project gezien. - Deze tabel bevat ook een startdatum en einddatum, dit is een belangrijk onderdeel om inzicht te krijgen op het tijdsverloop van projecten. De start- en einddatum zijn grotendeels aangegeven met 01-01-1900.
D_leverancier	- Kenmerken moeten hier aan toegevoegd of gelinkt worden.
D_datum_geleverd	- Sommige regels missen een administratie - De tabel bevat maar 3 administraties!
D_datum_besteld	- Sommige regels missen een administratie - De tabel bevat maar 3 administraties!
D_administratie	- Bevat B800 administraties, wat zijn de B800 administraties? - Worden deze B800 administraties nog gebruikt? - De 800 administraties halen geen data op
D_bestelorder	- Wat is "Dummy bestelregel" en welke toegevoegde waarde heeft deze kolom?
D_Bronmagazijn	- Naam reserveringsmagazijn en Reserveringsmagazijn bevatten geen data
D_doelmagazijn	- Naam reserveringsmagazijn en Reserveringsmagazijn bevatten geen data
D_Crediteur	- Zie opmerkingen in tabel op pagina 2
D_distributieroute	- Omschrijving bevat een code in sommige gevallen. Heeft dit relevant informatie en willen wij hier wat mee doen?
D_kostenplaats	- BI rayon en BI uitsluitingen zijn leeg, wat willen/kunnen wij met deze kolommen

	- Kostenplaats structuur implementeren vanuit CWD naar Acto (Dit moet in het DWH van CWD gebeuren)
D_persoon_aanvrager_adres, D_persoon_accordeerder_adres, D_persoon_inkoper_adres	- Deze drie tabellen zijn leeg. Waarom zijn deze leeg en willen we deze behouden ja/nee?
D_persoon_aanvrager_privacy, D_persoon_accordeerder_privacy, D_persoon_inkoper_privacy	- Deze drie tabellen bevatten twee kolommen die óók aanwezig zijn in: D_persoon_aanvrager, D_persoon_accordeerder & D_persoon_inkoper. Is deze tabel overbodig ja/nee?
D_vestiging	- B937 bevat afwijkende vestigingen - B939 bevat kostenplaatsen i.p.v. vestigingen op sommige regels. Dit komt door een verouderd bedrijfsproces waarbij kostenplaats als vestiging werd opgegeven - Uitsluitingen Vestiging kolom is leeg, is deze kolom overbodig? - Postcode is leeg, kan deze toegevoegd worden?

Tabel 16 data assessment opmerkingen

De tabel op de volgende pagina beschrijft de tabellen waarbij kolommen overbodig zijn en niet gebruikt worden, of waarvan er kolommen missen. Dit staat gespecificeerd in de opmerkingen.

Het is mogelijk dat er kolommen leegstaan, maar wel relevante data kunnen bevatten.

Het kan voorkomen dat het niet duidelijk is of extra kolommen (KvK & IBAN) toegevoegd moeten worden aan de tabel als of beter gelinkt kunnen worden.

Deze aanpassingen zijn een aanvulling op de al eerder besproken aanpassingen.

De prioriteit is als volgt geformuleerd:

1. Must have
2. Nice to have

Aanpassingentabel

Tabelnaam	kolommen	Opmerkingen	Prioriteit
D_vestiging	Postcode	Toevoegen van vestigings postcode	1
D_crediteuren	Allen	Locatie gegevens postcode, plaats en adres zijn leeg van: - Vestigingen - Factuur - Aflever Is er te achterhalen waarom deze niet gevuld worden met gegevens. Is het mogelijk om deze te vullen met de correcte data	1
D_crediteuren	KvK	KvK nummer toevoegen/koppelen	2
D_crediteuren	IBAN	IBAN nummer toevoegen/koppelen	2
D_persoon	Medewerker functie	Toevoegen	2

Tabel 17 data assessment aanpassingentabel

10.13 Bijlage 13 Datakwaliteit Meetinstrument



Data Driven Inkoop Croonwolver&dros

Data Management meetinstrument

AUTEUR	Dennis Loos
STUDENTNUMBER	S1096946
SCHOOL	Hogeschool Leiden
FACULTEIT	Informatica – Business Data & Management
STARTDATUM	06-02-2019
EINDDATUM	31-05-2019
ORGANISATIE	Croonwolver&dros
AFDELING	STAF-ICT-Informatie Management
VERSION NUMBER	3.0
VERSION DATE	07-06-2019
VERSION TYPE	Final

Versiemanagement

Versie	Datum	Status	Opmerkingen
0.1	28-02-2019	Ontwikkel	Opzet document
0.2	18-03-2019	Ontwikkel	Architectuur, Data Storage & Operations
0.3	20-03-2019	Ontwikkel	Data Security, Data Integration & Interoperability, Document & Content Management
0.4	21-03-2019	Ontwikkel	Data Governance, Data Quality, Meta Data, Data Warehousing & BI
0.5	22-03-2019	Ontwikkel	
0.6	27-03-2019	Ontwikkel	Data Quality aangepast
1.0	02-04-2019	Concept	Gevalideerd door de BI-manager
2.0	25-05-2019	Pre-Final	-
3.0	07-06-2019	Final	-



Contents

1. Aanleiding.....	101
Doel	101
Doelstelling.....	101
Doelgroep	101
Scope	102
2. Methode	103
3. Data Governance	106
4. Data Architectuur	108
5. Data Modeling & Design.....	109
6. Data storage & Operations.....	110
7. Data Security	111
8. Data integration & interoperability.....	113
9. Document & Content Management.....	114
10. Reference & Master Data.....	115
11. Data Warehousing & Business Intelligence.....	116
12. Metadata	117
13. Data Quality Management	118
Wat is datakwaliteit.....	118
Definitie datakwaliteit dimensies.....	119
Data Quality Management Croonwolder&dros	121
Activiteiten	121
Bibliography.....	125

Aanleiding

Ieder jaar vergroot de hoeveelheid gegenereerde data exponentieel. Dit komt mede door IoT en digitalisering. In 2020 wordt het 5G netwerk uitgerold, waardoor het genereren van data alleen nog maar meer wordt. Auto's kunnen met elkaar communiceren, Alle onderdelen in gebouwen kunnen verbonden worden met elkaar via een netwerk van sensoren, internet en andere technische apparaten.

Sinds 1 januari 2017 zijn Croon en Wolter & Dros gefuseerd tot Croonwolter&dros. Croon was gespecialiseerd in elektrotechniek en automatisering. Wolter & Dros was gespecialiseerd in werktuigbouwkunde en meet- en regeltechniek. De fusie betekende ook dat de bedrijfsprocessen, werkwijzen en de afspraken volgens een vaste norm moeten worden vastgelegd.

Hoge datakwaliteit is afhankelijk van structuur, precisie, procesmatig werken, de business en definities. Croonwolter&dros heeft veel verschillende administraties die data vastleggen in het datawarehouse. Met meerdere databronnen die de data invoeren groeit het risico op lage datakwaliteit. Maar wat is datakwaliteit nu precies? "datakwaliteit is de mate waarin data geschikt is voor het doel waarvoor ze gebruikt wordt".

Om te voorkomen dat de datakwaliteit verslechterd is er datakwaliteit management nodig. Met datakwaliteit management is een goede samenwerking tussen de business en IT nodig. Tools dragen zeker bij aan het verhogen van de datakwaliteit, maar de business moet ook het belang ervan inzien en hun bijdrage hiervan.

Op dit moment is er geen datakwaliteit management team binnen de organisatie. De data wordt ingevoerd en als er fouten gezien worden door gebruikers wordt dit (indien het belangrijk genoeg is om te melden) gemeld en opgelost door het BI-team.

Doel

Data management draagt bij aan het maken van beslissingen op verschillende niveaus binnen organisaties. Het doel van de data management meting om de informatiebehoeften van de organisatie en de belanghebbende te achterhalen. Een nulmeting om de huidige status van data management te toetsen, het waarborgen van de privacy en vertrouwelijkheid van gegevens van belanghebbenden, voorkomen van ongeoorloofde of ongepaste toegang, wijzigingen of gebruik van data en ervoor zorgen dat de data effectief gebruikt kan worden binnen de inkoopafdeling.

Doelstelling

De doelstelling is om een meetinstrument te ontwikkelen die de gesteldheid van data management binnen Croonwolter&dros kan aantonen. Het meetinstrument is zo geformuleerd dat dezelfde input altijd dezelfde output-score geeft. Met dit meetinstrument kan Croonwolter&dros ten allen tijden een meting doen en deze vergelijken met de voorafgaande meting(en).

Doelgroep

Het meetinstrument wordt ontwikkeld enkel voor de inkoopafdeling. Het moet mogelijk worden voor de data steward(s) om met de tool zelfstandig de datakwaliteit te beheren. De stuurgroep heeft belang bij dit product, om inzicht te krijgen in de data management van de organisatie. Het product wordt vanuit de Business IT afdeling geïmplementeerd. De BI-afdeling zal betrokken blijven bij het product en zal ondersteuning bieden bij vragen/opmerkingen.

Scope

Het complete DAMA DMBOK2 raamwerk wordt meegenomen in dit document. Binnen dit project worden enkele domeinen uitgevoerd en de overige domeinen worden beschreven met een advies over hoe en waarom dit belangrijk is voor Croonwolter&dros om het te implementeren.

Het meetinstrument toetst de gegevens van de Acto data marts. De betreffende data marts zijn *Inkoop* en *Crediteuren*. Naast de twee data marts zijn er een aantal externe Excel bestanden die meegenomen worden in het project, maar ook in de data kwaliteit meting. Het meetinstrument moet schaalbaar zijn en toe te passen zijn op variabele data marts/tabellen.

Methode

Het meetinstrument dat ontwikkeld wordt voor de inkoopafdeling zal een score geven op verschillende factoren. Deze metingen worden uitgevoerd op verschillende data management lagen binnen de inkoopafdeling. De methodiek dat gebruikt zal worden om de datakwaliteit te meten is het DAMA-DMBOK2 raamwerk.

Tijdens dit project worden niet alle onderdelen van het DAMA DMBOK2 raamwerk uitgevoerd. Binnen dit project worden de volgende gebieden gemeten:

Domein	Omschrijving	Aspecten	Uitvoeren?
Data Architectuur	De structuur van de data en dat gerelateerde bronnen als onderdeel van de bedrijfsarchitectuur.	Schaalbaarheid, Combineerbaar	Nee
Data Modeling & Design	Datamodelleren is het proces van het ontdekken, analyseren en scope van data vereisten, en vervolgens het weergeven en communiceren van deze gegevensvereisten in een datamodel.	Dit wordt aangeleverd.	Nee
Data Storage & Operations	Dataopslag en -operaties omvatten het ontwerp, de implementatie en ondersteuning van opgeslagen data om de waarde ervan tijdens de volledige levenscyclus, van creatie/verwerving tot verwijdering, te maximaliseren.	Data acquisitie behoud en afstemming van data	Nee
Data Security	Beveiliging van gegevens omvat de planning, ontwikkeling en uitvoering van beveiligingsbeleid en -procedures om correcte authenticatie, autorisatie, toegang tot en controle van gegevens- en informatieactiva te bieden. Het doel van gegevensbeveiliging is: informatiebronnen beschermen in overeenstemming met privacy- en vertrouwelijkheidsregels, contractuele overeenkomsten en bedrijfsvereisten.	Later afstemmen welke aspecten van Security binnen de scope liggen.	Ja
Data Integration & Interoperability	Data-integratie en interoperabiliteit (DII) beschrijft processen en activiteiten die zorgen voor de verplaatsing en beheren van gegevens binnen en tussen datastores, applicaties en organisaties. Data-interoperabiliteit is de mogelijkheid voor meerdere systemen om te communiceren.	Azure migratie en de Data marts van ACTO. Externe databronnen implementeren.	Nee
Document & Content Management	Document- en contentmanagement omvat het besturen van de vastlegging, opslag, toegang en gebruik van gegevens en informatie die zijn opgeslagen buiten het datawarehouse. De focus ligt op het handhaven van de integriteit van en het mogelijk maken van toegang tot documenten en andere ongestructureerde of semigestructureerde informatie waardoor deze ongeveer gelijkwaardig is aan gegevensbeheer voor relationele databases.	Excel bestanden	Nee

Reference & Master Data	Master Data Management (MDM) houdt controle in over master data en identificatiegegevens die een consistent gebruik van systemen mogelijk maken van de meest nauwkeurige en actuele gegevens over bedrijfsentiteiten. De doelen van MDM zijn onder meer het waarborgen van de beschikbaarheid van nauwkeurige, actuele waarden, terwijl de risico's die samen gaan met dubbele ID's worden verminderd. Reference Data Management (RDM) houdt controle in over gedefinieerde domeinwaarden en hun definities. Het doel van RDM is ervoor te zorgen dat de organisatie toegang heeft tot een complete set nauwkeurige en actuele waarden voor elk weergegeven concept.	De aspecten die getoetst worden van Dit domein vallen onder de Data Quality, Data Governance en Data Integration & Interoperability.	Nee
Data Warehousing & Business Intelligence	Data Warehousing beschrijft de operationele extractie, opschoning, transformatie, controle en laadprocessen die de gegevens in een datawarehouse onderhouden. Het data warehousing-proces richt zich op het mogelijk maken van een geïntegreerde en historische bedrijfscontext op operationele gegevens door het afdwingen van bedrijfsregels en het onderhouden van gepaste bedrijfsgegevensrelaties. Business Intelligence verwijst naar een reeks technologieën die dit soort data analyses ondersteunen. Beslissingsondersteunende tools en BI-tools maken query's, datamining, statistische analyse, rapportage, scenariomodellering, data-visualisatie en dashboarding mogelijk. Ze worden gebruikt voor alles, van budgettering tot geavanceerde analyses.	Later definiëren	Nee
Meta-Data	Metadata omvat informatie over technische- en bedrijfsprocessen, gegevensregels, en logische en fysieke gegevensstructuren. Het beschrijft de data zelf, de concepten die de data representeren en de verbindingen (relaties) tussen de data en concepten. Metadata helpt een organisatie haar gegevens, haar systemen en haar werkstromen te begrijpen.	-	Nee
Data Quality	De term gegevenskwaliteit verwijst zowel naar de kenmerken die zijn gekoppeld aan gegevens van hoge kwaliteit als naar de processen die worden gebruikt om de kwaliteit van gegevens te meten of te verbeteren.		Ja

	Gegevens zijn van hoge kwaliteit in die mate dat het voldoet aan de verwachtingen en behoeften van gegevensgebruikers.		
Data Governance	Data Governance is de uitoefening van autoriteit en controle (planning, monitoring en handhaving) op het beheren van 'data assets'. Data Governance gaat niet puur over data. Het vereist een combinatie van mensen, regels en procedures en ondersteunende technologieën om succesvol te zijn.	Naderhand specificeren.	Ja



Figuur 28 DAMA DMBOK2 raamwerk

Data Governance

Data Governance

Het doel van data governance is ervoor te zorgen dat gegevens correct worden beheerd, volgens beleid en best practices. Data governance focust zich op de manier waarop beslissingen worden genomen over gegevens en hoe mensen en processen met de gegevens omgaan. Voor Croonwolter&dros omvat een data governance-programma op dit moment:

- **Strategie:** Definiëren, communiceren, sturen en het centraal opslaan van de uitvoering van de strategie voor de data en databeheer.
- **Beleid:** Invoeren en draagvlak creëren van beleid met betrekking tot data en databeheer, toegang, gebruik, beveiliging en kwaliteit.
- **Normen en kwaliteit:** Invoeren en beheren van normen voor data kwaliteit en -architectuur.
- **Overzien:** Audits uitvoeren en ondersteuning bieden op het gebied van datakwaliteit, beleid en databeheer.
- **Naleving:** Zorgen dat de organisatie zich houdt aan wettelijke eisen omtrent data management.
- **Issue management:** Identificeren, definiëren en oplossen van problemen omtrent data security, -kwaliteit, -toegankelijkheid, -beheer en beleid en normen op een gestructureerde manier.
- **Projecten voor databeheer:** Ondersteunen van activiteiten en tools om databeheer te verbeteren.

Er dient een Data Governance programma opgezet moeten worden om deze doelen te bereiken. Dit programma ontwikkelt een beleid en procedures, data-stewards van verschillende niveaus en dienen de werknemers binnen de organisatie actief geïnformeerd te worden over de voordelen van verbeterde data governance.

Bedrijfsgebeurtenisgebieden moeten worden afgestemd op bedrijfsprocesbeheer aangezien elke gebeurtenisentiteit meestal overeenkomt met een bedrijfsproces. Data Architecture governance activiteiten omvatten:

- **Projectbegeleiding :** dit houdt in dat projecten moeten voldoen aan de vereiste Data Architecture-activiteiten, architecturale middelen moeten worden gebruikt en verbeterd en moeten worden geïmplementeerd volgens de vermelde architecturale normen.
- **Beheer van architectonische ontwerpen, levenscyclus en hulpmiddelen :** architecturale ontwerpen moeten worden gedefinieerd, geëvalueerd en onderhouden. Enterprise Data Architecture dient als een 'bestemmingsplan' voor lange termijnintegratie. Toekomstige staatsarchitectuur beïnvloedt projectdoelen en beïnvloedt de prioriteit van de projecten in het projectportfolio.
- **Definiëren van standaarden :** instellen van de regels, richtlijnen en specificaties voor hoe gegevens binnen de organisatie worden gebruikt.
- **Gegevens gerelateerde artefacten maken :** artefacten die naleving van governance-richtlijnen mogelijk maken.

Data stewards

Een data steward is een business professional die de verantwoording en verantwoordelijkheid voor gegevens en de processen ernaar zorgen voor een effectieve controle en gebruik van data assets. Data stewards kunnen geformuleerd worden als functietitels, maar het kan ook minder formeel geformuleerd worden. Bij grote organisaties is het vereist dat data stewards aangenomen en vastgelegd worden, zodat dit over de gehele organisatie duidelijk is.

Activiteiten van een data steward focust in de meeste gevallen op:

- **Maken en beheren van kernmetadata:** definitie en beheer van zakelijke terminologie, geldige gegevenswaarden en andere kritieke metagegevens. Stewards zijn vaak verantwoordelijk voor de Business woordenlijst van een organisatie, die het systeem voor het vastleggen van zakelijke termen met betrekking tot gegevens wordt.
- **Regels en normen documenteren:** definitie / documentatie van bedrijfsregels, gegevensstandaarden en regels voor gegevenskwaliteit. Stewards helpen deze regels onder de aandacht te brengen om ervoor te zorgen dat er binnen de organisatie consensus over is en dat ze consistent worden gebruikt.
- **Problemen met gegevenskwaliteit beheren : beheerders** zijn vaak betrokken bij het identificeren en oplossen van gegevens gerelateerde problemen of bij het faciliteren van het proces van oplossing.
- **Uitvoeren van operationele data governance-activiteiten :** de beheerders hebben de verantwoordelijkheid ervoor te zorgen dat het beleid en de initiatieven van dag tot dag en project-voor-project, data governance en initiatieven worden nageleefd. Ze moeten beslissingen beïnvloeden om ervoor te zorgen dat gegevens worden beheerd op manieren die de algehele doelen van de organisatie ondersteunen.

Afhankelijk van de complexiteit van de organisatie en de doelstellingen van haar DG-programma, kunnen formeel benoemde Data Stewards worden onderscheiden door hun plaats binnen een organisatie, door de focus van hun werk of door beide. Bijvoorbeeld:

- **Chief Data Stewards** kan data governance-organen voorzitten in plaats van de CDO of kan optreden als een CDO in een virtuele (op commissies gebaseerde) of gedistribueerde organisatie voor gegevensbeheer. Ze kunnen ook Executive Sponsors zijn.
- **Executive Data Stewards** zijn senior managers die zitting hebben in een Data Governance Council.
- **Enterprise Data Stewards** houden toezicht op een gegevensdomein in verschillende bedrijfsfuncties.
- **Business Data Stewards** zijn zakelijke professionals, meestal erkende materiedeskundigen, die verantwoordelijk zijn voor een subset van gegevens. Ze werken samen met belanghebbenden om gegevens te definiëren en te beheren.
- **Een eigenaar van gegevens** is een zakelijke gegevensbeheerder, die goedkeuringsbevoegdheid heeft voor beslissingen over gegevens binnen zijn domein.
- **Technische Data Stewards** zijn IT-professionals die opereren in een van de kennisgebieden, zoals Data Integration Specialists, Database Administrators, Business Intelligence Specialists, Data Quality Analysts of Metadata Administrators.
- **Coördinerende Data Stewards** leiden en vertegenwoordigen teams van zakelijke en technische Data Stewards in discussies tussen teams en met uitvoerende Data Stewards. Het coördineren van Data Stewards is met name belangrijk in grote organisaties.

Data Architectuur

Data architectuur is een essentieel onderdeel binnen data management. De data architectuur is de fundering van de data. De manier waarop de data architectuur is opgebouwd beïnvloed de functies, prestaties en de kosten van een systeem. Een goede data architectuur waarborgt een hoge datakwaliteit door juiste data modellen te gebruiken, relaties te leggen en door ervoor te zorgen dat de data schaalbaar is.

Het doel van data architectuur is om een verbinding te leggen tussen bedrijfsstrategieën en technologie.

De data architectuur van de afdeling inkoop is uitbesteed aan Acto. Acto moduleert de data van de Inkoopafdeling conform de dimensionaal modellering techniek van Kimball. De metrics die getoetst worden binnen de data architectuur zijn:

- **Schaalbaarheid** is het toetsen van de flexibiliteit van de data architectuur. Er wordt gekeken naar de impact van het uitbreiden en het implementeren van nieuwe databronnen op de huidige data.
- **Implementatietrends** is het toetsen op de mate waarop het aanpassen van de data architectuur invloed heeft op de werkzaamheden.
- **Modellen & Definities** kijkt naar de aanwezigheid van (technische) data modellen binnen de organisatie en de definities die binnen de organisatie geformuleerd zijn.
- **Kwaliteitsgericht** is de focus leggen op verbetering van de uitvoering binnen de ontwikkelingen van de organisatie. Het beheer van de data architectuur, definities en systemen wordt steeds ingewikkelder waardoor er risico's ontstaan.
- **Innovatiegericht** legt de focus op innovatieve kansen benutten om efficiënter de data aan te sturen, maar ook om te kunnen groeien als organisatie.



Data Modeling & Design

Het onderdeel data modeleren en design is niet relevant voor dit project. Dit wordt uitbesteed aan een andere organisatie. Aanpassingen aan het data model wordt in overleg met Croonwolver&dros en de leverancier van de data marts geïmplementeerd.

Data storage & Operations

Data storage & operations is het ontwerpen, implementeren en ondersteunen van opgeslagen gegevens om de waarde ervan tijdens de levenscyclus te maximaliseren. De twee hoofdactiviteiten op het gebied van gegevensverwerking en opslag zijn Database Technology Support en ondersteuning van databaseactiviteiten. Database Technology Support is specifiek voor het selecteren en onderhouden van de software die de gegevens opslaat en beheert. Ondersteuning van databasehandelingen is specifiek voor de gegevens en processen die de software beheert. Database Administrators spelen hierin een grote rol.

De doelstellingen van Data storage & Operations zijn:

- Beheer van beschikbaarheid van de data gedurende de levenscyclus
- Zorgen voor de integriteit van de data assets
- Beheer van de data operaties

Dit wordt omvat in twee domeinen:

- **Databaseondersteuning** is gericht op de activiteiten die de levenscyclus van de data. Vanaf het implementeren tot aan het back-uppen, opschonen en het wijzigen van de data. Ook de datawarehouse performance wordt gemonitord.
- **Ondersteuning van technologie** omsluit het definiëren van technische vereisten die vanuit de organisatie geformuleerd worden. Het installeren en beheren van technologie en het oplossen van problemen die betrekking hebben tot technologie.

Omgevingen

Een aspect van Operations is op het moment dat een wijziging doorgevoerd wordt binnen de organisatie, het getest is en te configureren met de huidige bedrijfsprocessen. Om dit te kunnen zijn er verschillende omgevingen om de nieuwe functionaliteiten te ontwerpen en te testen. Het is belangrijk dat de omgeving(en) veel op de productieomgeving lijkt, om zoveel mogelijk configuratiefouten te voorkomen.

Veerkracht en herstel

De veerkracht van een database toont aan hoe tolerant een systeem is tegen fouten. Als een systeem door kan blijven functioneren bij veel verwerkingsfouten, dan is het niveau van de veerkrachtigheid hoog. Wanneer het systeem crasht bij veelvoorkomende fouten, dan is de veerkrachtigheid van het systeem laag. Er zijn drie hersteltypes die een richtlijn geven over hoe snel een systeem hersteld wordt en waar het op focust:

- **Onmiddellijk herstel** zijn problemen die hersteld worden doormiddel van design. Hiervoor worden scripts ontwikkeld om de fouten te herstellen.
- **Kritiek herstel** is een plan om het systeem zo snel mogelijk te herstellen om vertragingen te minimaliseren.
- **Niet-kritisch herstel** zijn problemen die uitgesteld worden naar een ander tijdstip tot systeemfouten die kritischer zijn hersteld zijn.

Data Security

Beveiliging van data omvat het plannen, ontwikkelen en uitvoeren van procedures en beleid om de gegevens zo veilig als mogelijk te bewaren. Hierbij speelt de juiste authenticatie, autorisatie, toegang tot en controle van gegevens een grote rol. Bij Data security wordt gekeken naar welke gegevens beschermd moeten worden. Het doel van Data Security is informatiebronnen te beschermen in overeenstemming met de privacy- en vertrouwelijkheidsregels, contractuele overeenkomsten en bedrijfsovereenkomsten. Organisaties kunnen een boete krijgen voor het niet houden aan het reglement en contractuele verplichtingen. Datalekken kunnen reputatieschade, maar ook klantenschade veroorzaken. Contractuele verplichtingen of bedrijfsvereisten komen van verschillende afkomsten:

- **Stakeholders:** Organisaties moeten de privacy- en vertrouwelijkheidsvereisten van al hun stakeholders waarborgen
- **Overheid:** De organisatie dient de gegevens te beveiligen volgens bepaalde overheidsregels
- **Bedrijfseconomische belangen:** De gegevens van de organisatie, als deze gedeeld worden naar concurrenten, kan dit concurrentievoordeel opleveren.
- **Contractuele verplichtingen:** Contractuele overeenkomsten hebben ook invloed op vereisten voor gegevensbeveiliging.
- **Legitieme toegangsbehoeften:** Naast dat organisaties de gegevens moet beveiligen, dient het ook toegang te verlenen naar de juiste mensen of rollen de gegevens te gebruiken of onderhouden.

Op basis van de goederen en diensten van de organisatie worden de domeinen geselecteerd die relevant zijn. Data Security wordt op de volgende domeinen getest:

- **Beveiligingsprocessen (De vier A's):**
 - o **Access(toegang):** geef personen met een autorisatie toegang tot systemen.
 - o **Audit:** controleer acties en activiteiten om ervoor te zorgen dat de regels worden nageleefd en dat ze voldoen aan het beleid en de normen van de organisatie.
 - o **Autorisatie:** Personen rechten verlenen om toegang te krijgen tot specifieke gegevens of systemen. Dit kan gedaan worden op basis van rollen of per persoon.
 - o **Authenticatie:** De gebruikers valideren, wanneer een gebruiker zich probeert aan te melden in een systeem, moet het systeem een controle uitvoeren om te controleren of de persoon is wie hij/zij beweert is.
- **Wachtwoordnormen:** Welke maatregelen zijn er getroffen omtrent wachtwoorden om de veiligheid van de systemen te waarborgen. Denk hierbij aan complexiteit of de frequentie waarin een wachtwoord gewijzigd dient te worden.
- **Netwerkbeveiliging (Firewall, VPN):** Wat voor netwerkbeveiligingssysteem heeft de organisatie? (Alleen een firewall is niet voldoende, een goede beveiliging bestaat uit meerdere lagen. Des te meer lagen een beveiligingssysteem heeft, des te veiliger het is. Er moet rekening gehouden worden met de gegevens die beveiligd worden. Gevoelige data moet meer beveiligingslagen hebben dan niet-gevoelige data. Wanneer je een drie stappen verificatie moet invoeren om toegang te krijgen tot informatie wat er als lunch geserveerd wordt vandaag in de kantine, dat is overbodig goed beveiligd.
- **Systeembeveiliging:** Privileges, gedeelde accounts
- **malware (virus, worm, trojan horse, adware)**
- **Beveiligingsbeleid** wordt onderverdeeld in drie verschillende aspecten:



- Bedrijf beveiligingsbeleid: algemeen beleid voor toegang van werknemers tot middelen, standaarden en toegangsniveaus voor beveiliging op basis van positie binnen de organisatie.
- IT-beveiligingsbeleid: wachtwoordbeleid, identiteitsbeheer en normen voor structuren
- Gegevensbeveiligingsbeleid: Databaserollen, gebruikersgroepen en informatiegevoeligheid.

Data integration & interoperability

Data Integration & Interoperability (DII) beschrijft de processen die de verplaatsing en de consolidatie van gegevens binnen datastores, applicaties en organisaties. Bij de integratie van gegevens behoudt de data zijn waarde. Data-interoperability is de mogelijkheid voor meerdere systemen om te communiceren.

Het doel van DII is om gegevens beschikbaar te maken in de juiste format voor de gebruikers. DII ondersteund Business Intelligence, omdat dit gericht is op het transformeren en het integreren van data uit bronnen naar data marts die gebruikt worden om de data te visualiseren. De centrale gebieden van DII is het Extract, Transform & Load (ETL) proces.

Onderdelen die getoetst worden die relevant zijn voor de organisatie:

- **Archivering:** gegevens die niet vaak gebruikt worden of niet actief worden gebruikt, kunnen worden verplaatst naar een alternatieve database. Hierin worden gegevens geplaatst niet (bijna) niet meer gebruikt worden binnen een organisatie.
- **ETL tools** bevatten een ontwerp van de activiteiten en ondersteunen bij de acties die ondernomen moeten worden.
- **Data afstemmen** is nodig om de gegevens tussen de verschillende divisies of administraties binnen de organisatie te integreren.
- De **data volume, beschikbaarheid** en **snelheid** wordt in het domein Data Quality Management getoetst.

Document & Content Management

Document & contentmanagement is het besturen van de definities, opslag en de toegang van gegevens die zijn opgeslagen buiten het datawarehouse. De nadruk ligt op het waarborgen van de integriteit en het mogelijk maken om de documenten te kunnen combineren met gegevens uit het datawarehouse.

Het doel van het implementeren van best practices rond document- en contentbeheer zijn:

- Effectief en efficiënt ophalen en gebruiken van gegevens en informatie in ongestructureerde data bestanden.
- Zorgen voor integratiemogelijkheden tussen gestructureerde data en ongestructureerde data.
- Voldoen aan wettelijke verplichtingen en de verwachtingen van de klant.
- Iedere gebruiker die de bestanden kan aanpassen werken volgens vastgestelde procedures en beleidsregels. Deze gebruikers moeten worden betrokken bij het ontwikkelen van het beleid en de regels omtrent het documentenbeheer.

De domeinen van Document & Content Management die in het meetinstrument getoetst worden zijn:

- Meta-data: wordt gebruikt om inhoud te inventariseren, de meta-data is nodig om de context van de inhoud te begrijpen. Lage kwaliteit/slechte meta-data kan ervoor zorgen dat de inhoud verkeerd of niet begrepen wordt door de gebruikers.
- Documentbeheer & versiebeheer omvat de processen, technieken en technologieën voor het beheren en organiseren van documenten en records gedurende hun levenscyclus. Documentbeheersystemen bevatten ook versiebeheer om bij te houden wie, wanneer, wat heeft bijgewerkt.

Dit volwassenheidsmodel beschrijft de kenmerken van de informatie governance- en archiveringsomgeving op vijf niveaus van volwassenheid voor elk van de acht GARP-principes:

- Sub niveau van niveau 1 : Informatiebeheer en archiveringsproblemen komen niet of slechts minimaal aan bod
- Niveau 2 in ontwikkeling : erkenning ontwikkelen dat informatiebeheer en archivering van invloed kunnen zijn op de organisatie
- Level 3 Essentieel : Minimumeisen waaraan moet worden voldaan om te voldoen aan de wettelijke en reglementaire vereisten
- Niveau 4 Proactief : er is een proactief informatiebeheerprogramma opgezet met de focus op continue verbetering
- Level 5 Transformationeel : Information governance is geïntegreerd in de bedrijfsinfrastructuur en bedrijfsprocessen

Reference & Master Data

Een organisatie ter grootte van Croonwolver&dros beschikt over verschillende bedrijfseenheden. Deze bedrijfseenheden willen toegang hebben tot dezelfde klantenlijsten, leveringsopties, raamcontracten of API's. De gebruikers die met gegevens werken, gaan ervan uit dat er een consistent niveau is in de gehele organisatie. Reference & Master Data is het managen van gedeelde data om organisatorische doelen te bereiken, risico's omtrent redundante data te verlagen, hoge datakwaliteit te waarborgen, en de kosten van data integratie te verlagen.

Referentiegegevens zijn beschrijvingstabellen die worden gebruikt om de andere gegevens te karakteriseren of te classificeren. Voorbeelden van referentiegegevens zijn: Meeteenheden, administratiecodes of productgroepen.

Master Data zijn gegevens voor de bedrijfsobjecten (bijvoorbeeld werknemers, klanten, producten of financiële structuren) die context bieden voor zakelijke transacties en analyses. Master Data wordt in de gehele organisatie gebruikt, zoals leveranciers bij Croonwolver&dros.

Het belang van Reference & Master Data bij een organisatie:

- Het mogelijk maken van data assets tussen bedrijfsdomeinen en applicaties binnen de organisatie te delen
- Kosten en complexiteit te verlagen door het standaardiseren van data modellen en integratie

De elementen waar Reference Data Management & Master Data Management op beoordeeld wordt met dit meetinstrument zijn:

- Consistente gegevenssets: De verschillende bedrijfsdomeinen beschikken over dezelfde complete, actuele en consistente data.
- Gedeelde gegevens: De referentie en de master data zijn zo beheerd dat het beschikbaar is voor de hele organisatie
- Verantwoordelijkheid: De eigenaarschap of verantwoordelijkheid voor de definitie en het beheer.
- Standaardisatie zorgt ervoor dat de data voldoet aan de standaard formaten of velden zoals telefoonnummer of postcodes.

Data Warehousing & Business Intelligence

De belangrijkste driver van data warehousing is het ondersteunen van operationele functies en Business Intelligence (BI) activiteiten. Data warehousing operaties zijn het ophalen, opschonen, transformeren, controleren en wegschrijven van gegevens in een datawarehouse omgeving. Het proces van datawarehousing heeft als doel, het mogelijk maken van verschillende soorten data in één systeem zodat er beslissingen genomen kunnen worden op basis van data met rapportages en analyses. BI is het analyseren van gegevens en deze om te zetten in informatie, dit leidt vervolgens tot kennis en hiermee kan de organisatie actie ondernemen om de organisatie slimmer en efficiënter te laten werken.

Het Datawarehouse van de inkoop- en crediteuren afdeling is gemoduleerd volgens Kimball's dimensioneel model, met dimensies en feiten.

Metadata

Metadata helpt om de gegevens van een organisatie te begrijpen. Het is een onderdeel van de datakwaliteit uit en van het database beheer. Metadata kan ook omschreven worden als Data over data. Zonder betrouwbare metadata weet een organisatie niet welke data zij hebben, wat het inhoudt en waar het vandaan komt en is het onmogelijk om het te beheren. (Metadata kan gekeurd worden volgens de ISO / IEC 11179 registratienorm voor metadata).

Het doel van metadata is:

- Verzamelen en integreren van metadata uit verschillende bronnen, zodat gebruikers de overeenkomsten en verschillen tussen de gegevens uit verschillende delen van de organisatie begrijpen
- De organisatie voorzien van een standaard manier om data te omschrijven
- Waarborgen van datakwaliteit en toegankelijkheid door correcte context

Metadata wordt in drie categorieën ingedeeld:

- Bedrijfs-metadata richt zich op de inhoud en de staat van de gegevens. Voorbeelden van bedrijfs-metadata zijn; definities en beschrijvingen van datasets, tabellen en kolommen, Datastandaarden of datamodellen.
- Technische metadata beschrijven de technische eigenschappen van een component zoals de bestandsgrootte, het type bestand of de kolomeigenschappen.
- Operationele metadata beschrijft details van de verwerking en toegang tot de data. Bijvoorbeeld logboeken, resultaten van audits of uitvoeringstijden van query's.

De metadata wordt op verschillende onderdelen getoetst. Procedures, definities, Gebruik en vertegenwoordiging/beheer zijn speerpunten die belangrijk zijn voor de kwaliteit van data management.

Data Quality Management

Data Quality Management is het plannen, implementeren en controleren van activiteiten die de datakwaliteit managet om de wensen en de eisen van de gebruikers te voldoen. De drijfveren van data quality management is het vergroten van de waarde van de organisatie haar data en de mogelijkheden om het te gebruiken, het verminderen van risico's en kosten in verband met slechte datakwaliteit, en het verhogen van de efficiëntie en productiviteit van de organisatie.

De internationale standaard voor datakwaliteit is de ISO 8000, de ISO definieert deze norm als: "De mogelijkheid om gegevens te creëren, verzamelen, opslaan, onderhouden, overdragen, verwerken en presenteren om bedrijfsprocessen tijdig en kosteneffectief te ondersteunen vereist zowel een goed begrip van de kenmerken van de gegevens die de kwaliteit ervan bepalen, en een vermogen om datakwaliteit te meten, beheren en te rapporteren.").

Wat is datakwaliteit

Datakwaliteit is data die voldoet aan de eisen die een organisatie opgeeft. De eisen zijn afhankelijk per organisatie. Iedere organisatie heeft zijn eigen belang bij datakwaliteit.

Binnen datakwaliteit gaat het meer dan de aanwezigheid van fouten. Datakwaliteit kan in zes verschillende aspecten onderverdeelt worden zoals in onderstaande Tabel 18 Datakwaliteit dimensies is weergegeven.

Dimensie	Omschrijving
Tijdigheid	Dit is de mate waarin data op tijd aanwezig is.
Volledigheid	Aanwezigheid van de volledige verwachte dataset.
Accuraatheid	Dit geeft aan in hoeverre de data de werkelijkheid weergeeft.
Consistentie	Wanneer de consistentie optimaal is, heeft de gehele organisatie dezelfde coderingen en verwijzingen voor de data.
Integriteit	Dit is de moeilijkheidsgraad van de data, het komt veel voor dat cijfers in een database moeilijk te begrijpen zijn of dat de definitie onbekend is, om dit te voorkomen is een goede documentatie en meta-data vereist.
Uniekheid	De mate waarin data eenmalig in de dataset bestaat en uniek identificeerbaar is.
Beschikbaarheid	De mate waarin de data beschikbaar is wanneer de gebruiker het nodig heeft.

Tabel 18 Datakwaliteit dimensies

Definitie datakwaliteit dimensies

In onderstaande tabellen staan de dimensies omschreven, inclusief de meeteenheid en relevante informatie om meer duidelijkheid te geven aan de taak per dimensie.

Titel	Tijdigheid (Timeliness)
Definitie	De mate waarin data op tijd aanwezig is.
Graadmeter	Verschil in tijd
Meeteenheid	Tijd
Type meting (Assessment, Continue, Discreet)	Assessment en Continue
Gerelateerde dimensies	Accuraatheid, omdat het met de tijd vervalst.
Voorbeeld	Het systeem dat iedere nacht de data ophaalt heeft een error opgelopen. De data wordt in de nacht niet opgehaald, waardoor de dag erop de datawarehouse geen data bevat van die dag ervoor.

Tabel 19Tijdigheid

Titel	Volledigheid (Completeness)
Definitie	De mate waarin de volledige verwachte dataset aanwezig moet zijn.
Graadmeter	Een maat voor de afwezigheid van lege, of nietszeggende waarden.
Meeteenheid	Procent
Type meting (Assessment, Continue, Discreet)	Assessment
Gerelateerde dimensies	Integriteit en Accuraatheid
Voorbeeld	Bij een organisatie worden medewerkers gevraagd om een telefoonnummer op te geven voor noodgevallen. De organisatie heeft 1000 medewerkers, hiervan hebben er 850 een telefoonnummer voor noodgevallen ingevoerd. Er is een analyse uitgevoerd op de Noodgevallen telefoonnummers gegevens. De bevinding hieruit is dat $850/1000 \cdot 100 = 85\%$ van de dataset volledig is.

Tabel 20 Volledigheid

Titel	Accuraatheid (Accuracy)
Definitie	De mate waarin de data in overeenstemming is met de werkelijkheid.
Graadmeter	De mate waarin de gegevens overeenkomen met de kenmerken van de werkelijke objecten.
Meeteenheid	Percentage
Type meting (Assessment, Continue, Discreet)	Assessment, Continue
Gerelateerde dimensies	Integriteit
Voorbeeld	Een Nederlandse organisatie is gefuseerd met een Amerikaanse organisatie. Met deze fusie worden de normen van de Amerikaanse organisatie aangehouden. De datum wordt in Nederland als DD/MM/YYYY genoteerd, maar in Amerika als MM/DD/YYYY.

Tabel 21 Accuraatheid

Titel	Consistentie (Consistency)
Definitie	De mogelijkheid om data op een juiste manier over verschillende platforms te gebruiken.
Graadmeter	Data vergelijken over verschillende systemen
Meeteenheid	Percentage
Type meting (Assessment, Continue, Discreet)	Assessment en Discreet
Gerelateerde dimensies	Accuraatheid, Uniekheid en integriteit
Voorbeeld	Een student zijn geboortedatum is hetzelfde geformuleerd en bevat dezelfde data in de schoolregistratie als in de studenten database.

Tabel 22 Consistentie

Titel	Integriteit (Validity)
Definitie	De mate waarin de juistheid en de consistentie van de data behouden wordt. De data kan niet onbedoeld aangepast worden.
Graadmeter	Vergelijking tussen de data en de meta-data.
Meeteenheid	Percentage
Type meting (Assessment, Continue, Discreet)	Assessment, Continue en Discreet
Gerelateerde dimensies	Accuraatheid, Compleetheid, consistentie en uniekheid

Tabel 23 Integriteit

Titel	Uniekheid (uniqueness)
Definitie	De mate waarin er sprake is van dubbele data binnen een dataset.
Graadmeter	Een analyse van het aantal werkelijke dingen versus het aantal records van dingen in de dataset.
Meeteenheid	Percentage
Type meting (Assessment, Continue, Discreet)	Discreet
Gerelateerde dimensies	Consistentie
Voorbeeld	Een school heeft 380 studenten ingeschreven staan en 2401 studenten in totaal ingeschreven staan die ooit op die school hebben gestudeerd. De database heeft 3002 studenten in de database staan. Dit geeft een uniekheid waarde van $2701/3002 * 100 = 89,97\%$.

Tabel 24 Uniekheid

Titel	Beschikbaarheid (Availability)
Definitie	De mate waarin de data beschikbaar is voor gebruikers.
Graadmeter	Het aantal keren dat de data beschikbaar is als het wordt gebruikt.
Meeteenheid	Percentage
Type meting (Assessment, Continue, Discreet)	Discreet
Gerelateerde dimensies	Tijdigheid, volledigheid
Voorbeeld	Een systeembeheerder heeft data nodig over de werknemers van de organisatie. Deze data staat in de database. In 10 dagen haalt hij één keer per dag de data op. 8 van de 10 keer krijgt hij alle werknemers te zien en de andere 2 keer missen er regels. In 80% van de gevallen is de data beschikbaar.

Tabel 25 Beschikbaarheid

Data Quality Management Croonwolter&dros

De kern van dit meetinstrument is de datakwaliteit. In dit domein wordt er verder ingezoomd binnen het meetinstrument. De onderdelen die getest gaan worden, worden op basis van de projecteisen geformuleerd. Het doel van datakwaliteit management is;

- Het definiëren van standaarden en specificaties voor datakwaliteitscontroles
- Het ontwikkelen van een gestandaardiseerde aanpak om gegevens toepasbaar te maken voor de gebruikers.
- Definiëren en implementeren van processen voor het meten, waarborgen en rapporteren van het datakwaliteitsniveau,
- Innoveren voor mogelijkheden om de datakwaliteit te verhogen, waarmee de organisatie nieuwe kennis en inzichten kan verkrijgen.

Activiteiten

Om de datakwaliteit te toetsen worden er procesmatig activiteiten doorlopen. Met het gebruik van deze methode is de risico dat er onderdelen over het hoofd worden gezien of de kans dat er fouten gemaakt worden geminimaliseerd.

Definieer data van hoge kwaliteit

Om vage definities zoals “De data moet kloppen” of “we willen hoge kwalitatieve data”, worden er definities en prioriteiten. Om de huidige status en de doelstellingen te achterhalen, worden er vragen opgesteld:

- Wat wordt er bedoeld met “hoge datakwaliteit”?
- Wat is de impact van lage kwalitatieve data op bedrijfsactiviteiten en strategie?
- Hoe kan data van hoge kwaliteit bijdragen aan de bedrijfsstrategie?
- Welke data heeft de hoogste prioriteit voor de organisatie?
 - o Welke data entiteiten zijn van groot belang dat ze kloppen met de werkelijkheid voor het eindresultaat?
 - o Welke geldigheidsregels zijn er op veldniveau?

- Welke prioriteiten heeft het verhogen van de datakwaliteit?
- Welke pijnpunten ondervindt de afdeling op dit moment?
- Wat is de tolerantie voor lage datakwaliteit?
- Welke data governance elementen zijn er nodig om de datakwaliteit te managen?

Bovenstaande vragen worden afgenomen van de stakeholders die antwoord kunnen geven op deze vraagstukken.

Datakwaliteit strategie

Op basis van de bevindingen van de vragen en de bedrijfsvereisten, worden er prioriteiten gesteld en de datakwaliteitsnormen gedefinieerd. De methode moet rekening houden met de werknemers die het instrument in de toekomst gaan gebruiken. De afname van de metingen houdt rekening met de datakwaliteitsnormen en stelt het instrument daarop af. Het is niet mogelijk om alle regels en manieren de eerste beoordeling aan te pakken. Het gaat om het proces en het begrijpen van de regels waarmee de organisatie te maken heeft, het achterhalen en definiëren van de regels is een continue proces.

Het ontwikkelen van een raamwerk en methodologie zal zowel de strategie als de tactiek helpen en tegelijk de voortgang en impact te meten. Een kader kan methoden omvatten om:

- Begrijp en prioriteer zakelijke behoeften
- Identificeer de gegevens die cruciaal zijn voor het voldoen aan bedrijfsbehoeften
- Definieer bedrijfsregels en datakwaliteitsnormen op basis van bedrijfseisen
- Deel bevindingen en krijg feedback van belanghebbenden
- Prioriteiten stellen en problemen beheren
- Identificeer en prioriteer mogelijkheden voor verbetering
- Meet, bewaak en rapporteer over de gegevenskwaliteit
- Beheer metagegevens die zijn geproduceerd via processen voor gegevenskwaliteit
- Integreer datakwaliteitscontroles in zakelijke en technische processen

Eerste datakwaliteit beoordeling

Zodra de kritieke behoeften en de data die hiervoor gebruikt worden bekend zijn, is het tijd om de data te analyseren. Hierbij wordt de inhoud, de relaties en het tegenover de regels en de verwachtingen van de organisatie leggen. Het doel van deze beoordeling is om meer te weten te komen over de data om een actieplan voor verbetering te ontwikkelen. Hierbij zullen de volgende stappen ondernomen worden:

- De doelen definiëren van de beoordeling
- Benoemen van de te beoordelen data, hiervoor wordt een kleine dataset gebruikt
- Benoemen van het gebruik van de data
- Risico's identificeren, inclusief de impact
- Data controleren op basis van de gestelde regels en normen
- De bevindingen worden gebruikt voor de volgende stappen

Gebruik bevindingen als basis voor planning van:

- Het herstellen van data fouten of issues, idealiter naar hun diepere oorzaken
- Controles en procesverbeteringen om te voorkomen dat problemen zich herhalen
- Continue controle en rapportage

Verbeteringen identificeren en prioriteren

De kennis die vergaard is tijdens de eerste datakwaliteit beoordeling wordt doorgenomen. Op basis van de bevindingen worden verbeteringen geïdentificeerd, maar ook geprioriteerd. Sommige problemen worden op korte termijn aangepakt. De strategische veranderingen worden geïdentificeerd om aan te pakken op de lange termijn.

Er moet rekening gehouden worden bij het implementeren van nieuwe procedures, systemen en activiteiten. (Oude) Systemen, bedrijfscultuur, oude gegevens en complexiteit van de data architectuur kunnen complicaties geven.

Bij het opstellen van verbeteringen moet er een zakelijke impact zijn, er moet een positief rendement uitkomen bij het investeren van de verbeteringen op de data. Hoe hoger het rendement, des te belangrijker het is om de verbeteringen te implementeren.

Doelen definiëren voor verbetering van datakwaliteit

De nieuwe kennis die is opgedaan tijdens de vorige beoordeling worden meegenomen naar de volgende beoordeling. Verbeteringen kunnen op verschillende niveaus worden gedefinieerd. Het kunnen quick wins zijn zoals correcties op foute velden, tot aan het herstel van de oorzaak. Problemen kunnen snel en goedkoop opgelost worden, maar er kunnen ook strategische veranderingen aangepakt moeten worden op langere termijn en/of hoge kosten.

Bij het definiëren van verbeteringen dient er rekening gehouden te worden met systeembependingen, oude data, data architectuur of bedrijfscultuur. Stel daarom de doelen zo SMART als mogelijk op, om te voorkomen dat de genoemde beperkingen de veranderingen blokkeren.

Het doel moet een zakelijke impact hebben, anders heeft het geen toegevoegde waarde om deze te verbeteren. Bijvoorbeeld: De inkoper krijgt alle informatie van een leverancier waarmee hij/zij mee gaat onderhandelen voor prijsafspraken de volgende dag.

Datakwaliteit modificaties ontwikkelen en implementeren

- **Problemen met gegevens-invoer-interfaces** : Slecht ontworpen interfaces voor gegevensinvoer kunnen bijdragen aan problemen met de gegevenskwaliteit. Als een interface voor gegevensinvoer geen bewerkingen of besturingselementen bevat om te voorkomen dat onjuiste gegevens in de systeem gegevens worden geplaatst, nemen de processors waarschijnlijk snelkoppelingen, zoals het overslaan van niet-verplichte velden en het niet bijwerken van standaardvelden.
- **Plaatsing van lijstitems** : zelfs eenvoudige functies van gegevens-invoer-interfaces, zoals de volgorde van waarden in een vervolgkeuzelijst, kunnen bijdragen aan fouten bij het invoeren van gegevens.
- **Overbelasting van velden**: sommige organisaties hergebruiken velden in de tijd voor verschillende zakelijke doeleinden in plaats van wijzigingen in het gegevensmodel en de gebruikersinterface aan te brengen. Deze praktijk resulteert in een inconsistente en verwarrende populatie van de velden.
- **Trainingsproblemen** : Gebrek aan proceskennis kan leiden tot onjuiste gegevensinvoer, zelfs als er besturingselementen en bewerkingen aanwezig zijn. Als gegevensverwerkers zich niet bewust zijn van de gevolgen van onjuiste gegevens of als ze eerder geneigd zijn tot snelheid dan nauwkeurigheid, zullen ze waarschijnlijk keuzes maken op basis van andere stuurprogramma's dan de kwaliteit van de gegevens.

Bibliography

- Acto. (2019, 05 27). *Acto*. Opgehaald van Acto: <https://www.acto.nl>
- arbocatalogus, d. (2019, 02 14). *arbocatalogus*. Opgehaald van deArbocatalogus: <https://www.dearbocatalogus.nl/handreiking-intern-transport-en-logistiek/52-beoordelingstool-2-risicobeoordeling-van-fine-en-kinney>
- Brautigam, I. (2019, 02 19). *Scrumguide*. Opgehaald van Scrumguide: <https://scrumguide.nl/sprint-review/>
- DAMA. (2017). *DAMA DMBOK2*. -: DAMA.
- DataLadder. (2019, 05 25). *DataLadder*. Opgehaald van DataLadder: <https://www.Dataladder.com>
- ERP. (2019, 02 21). *ERPdefinitions*. Opgehaald van ERPdefinitions: <https://www.erpsystemen.nl/wat-is-erp>
- Fred Conijn, J. d. (2018, 01 01). *cmweb*. Opgehaald van cmweb: <https://cmweb.nl/2018/01/datamanagement-kwaliteit-gegevens-bepaalt-kwaliteit-informatie/>
- Jonker, R. (2019, 02 28). *Compact*. Opgehaald van Compact: <https://www.compact.nl/articles/datakwaliteitsonderzoek/>
- Kimball. (2019, 05 20). *KimballGRoup*. Opgehaald van KimballGroup: <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/>
- Leansixsigma. (2019, 02 06). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/overzicht-stakeholders>
- Leansixsigma. (2019, 05 07). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/wat-zijn-key-performance-indactors>
- Microsoft. (2019, 4 20). *azure.microsoft*. Opgehaald van azure.microsoft: <https://azure.microsoft.com/nl-nl/overview/what-is-devops/>
- Motion10. (2019, 02 06). *Motion10*. Opgehaald van Motion10: <https://www.motion10.nl/>
- Oreilly. (2019, 03 14). *Oreilly*. Opgehaald van Oreilly: <https://www.oreilly.com/topics/oreilly-learning>
- Redactie. (2012, 10 26). *Business analytics*. Opgehaald van Business analytics: <https://www.business-analytics.biz/blogs/de-waarde-van-data-governance-en-het-zichtbaar-maken-ervan/>
- redactie, C. (2019, maart 22). *Computerworld*. Opgehaald van Computerworld: <https://computerworld.nl/business-intelligence/61874-wat-is-datawarehousing>
- Swaen, B. (2019, 04 20). *Scribbr*. Opgehaald van Scribbr: <https://www.scribbr.nl/onderzoeksmethoden/kwalitatief-vs-kwantitatief-onderzoek/>
- Tozzi, C. (2019, 02 28). *Syncsort*. Opgehaald van Syncsort: <https://blog.syncsort.com/2018/02/data-quality/how-to-measure-data-quality-7-metrics/>
- University, S. (2019, 05 09). *Stanford*. Opgehaald van Stanford: <https://web.stanford.edu/dept/pres-provost/cgi-bin/dg/wordpress/wp-content/uploads/2012/09/DG-Data-Definition-Best-Practices.docx>



- Verheul, M. (2019, 05 16). *Agilescrumgroup*. Opgehaald van Agilescrumgroup:
<https://agilescrumgroup.nl/welke-rollen-zijn-er-binnen-scrum/>
- Wethinknext. (2019, 02 07). *wethinknext*. Opgehaald van wethinknext:
<https://wethinknext.com/stakeholder-mapping/>
- Wikipedia. (2019, 03 07). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Databeveiliging>
- Wikipedia. (2019, maart 19). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Data-architectuur>

10.14 Bijlage 14 Plan van Aanpak



Data Driven Inkoop Croonwolver&dros

Plan van Aanpak

AUTEUR	Dennis Loos
STUDENTNUMBER	S1096946
SCHOOL	Hogeschool Leiden
FACULTEIT	Informatica – Business Data & Management
STARTDATUM	06-02-2019
EINDDATUM	31-05-2019
ORGANISATIE	Croonwolver&dros
AFDELING	STAF-ICT-Informatie Management
VERSION NUMBER	3.2
VERSION DATE	23-05-2019
VERSION TYPE	Final

Samenvatting

Dit document beschrijft het plan van Croonwolt&Dros om een oplossing te ontwikkelen voor het visualiseren van de huidige bedrijfsprocessen binnen de inkoopafdeling. De organisatie, de afstudeerder en de school worden beschreven in het hoofdstuk “De organisatie” met de samenhangende relaties en begeleiding. Vervolgens wordt er een stakeholderanalyse gedaan om de belanghebbende te verzamelen, met de impact die zij op het project hebben. Het hoofdstuk projectdoel beschrijft wat de organisatie wil bereiken met het implementeren van dit project. De probleemstelling omschrijft het probleem waar de organisatie mee kampt. Vervolgens worden de afhankelijkheden, de middelen en de uitsluitingen beschreven om de scope van het project aan te geven. De deliverables zijn de producten die opgeleverd worden, met deelproducten en een Product Breakdown Structure. Het hoofdstuk projectaanpak omschrijft de wijze waarop er gewerkt wordt om de oplossing te realiseren. Vervolgens is er een projectplanning gemaakt en tenslotte is er een risicoanalyse gedaan op het project.

De volgende hoofdstukken omschrijven de kern van het project:

- Projectdoel
- Probleemstelling
- Projectaanpak
- Projectplanning

Begrippenlijst

Woord	Betekenis
Raamcontract	Een afspraak tussen een leverancier en een klant om een bepaald product of dienst voor een vaste looptijd voor een vooraf afgesproken prijs, en tegen vooraf afgesproken voorwaarden te leveren.
Product Breakdown Structure	Een overzicht met alle producten die opgeleverd worden voor het project.
Enterprise Resource Planning	Een systeem dat de productiviteit van een organisatie optimaliseert door automatische afhandelingen van logistieke, administratieve en financiële bedrijfsprocessen onder te brengen in één managementsysteem.
Real-time	Een taak in een door software gecontroleerd systeem wordt realtime genoemd als de gecombineerde reactie- en uitvoertijd van de taak korter is dan de maximale tijd die is toegestaan, rekening houdend met invloeden van buitenaf.
KPI	Kritieke prestatie-indicatoren, afgekort KPI's, zijn variabelen om prestaties van ondernemingen te analyseren.
Data mart	Een data mart is een verzameling van gegevens, vergelijkbaar met een datawarehouse, maar meestal met een kleinere hoeveelheid aan gegevens en vaak ingericht voor een specifiek doel.

Tabel 26 begrippenlijst

Versiebeheer

Versie	Datum	Status	Opmerkingen
0.1	06-02-2019	Ontwikkel	Opzet document, Organisatie beschrijving
0.2	08-02-2019	Ontwikkel	Projectdoel en Probleemstelling opgezet
0.3	11-02-2019	Ontwikkel	Projectaannames en deliverables aangepast
0.4	13-02-2019	Ontwikkel	Deliverables ontworpen
0.5	14-02-2019	Ontwikkel	Stakeholders aangepast, risicoanalyse, probleemstelling aangepast
0.6	15-02-2019	Ontwikkel	Deliverables, PBS, Projectaanpak, Planning afgerond.
1.0	19-02-2019	Concept	Feedback na PvA bespreking verwerkt, titelpagina vormgegeven
1.1	19-02-2019	Concept	Versie gewijzigd naar Concept versie. Samenvatting geschreven.
1.2	20-02-2019	Concept	Feedback van Rob Kabel doorvoeren.
2.0	21-02-2019	Pre-Final	Omgezet naar Pre-Final
3.0	07-03-2019	Final	Feedback verwerkt en omgezet naar Final versie
3.1	18-02-2019	Final	Feedback verwerkt.
3.2	23-05-2019	Final	Spellingscontrole.

Inhoudsopgave

Samenvatting.....	128
Begrippenlijst.....	128
Versiebeheer	129
1 Achtergrond.....	131
2 De organisatie.....	132
2.1 Gegevens	15
2.2 Begeleiders	16
2.3 Mede-afstudeerder	17
2.4 Vorm van begeleiding.....	134
2.5 Organigram van de organisatie	134
2.6 Stakeholders	136
3 Projectdoel	137
3.1 Bedrijfsdoelstelling	137
4 Probleemstelling.....	138
5 Projectaannames.....	139
5.1 Uitsluitingen	139
5.2 Afhankelijkheden.....	139
5.3 Middelen	140
6 Deliverables	141
6.1 Productbeschrijvingen.....	141
6.2 Product Breakdown Structure	143
7 Projectaanpak.....	144
7.1 B-competenties	145
7.2 A-Competenties.....	146
8 Projectplanning	147
9 Risicoanalyse	149
9.1 Risicowaarden	151
10 Oplevering	152
11 Bronvermelding.....	153

1 Achtergrond

Croonwolter&dros loopt al meer dan 140 jaar voorop bij nieuwe ontwikkelingen. Ook in deze tijd, waarin de rol van techniek steeds belangrijker wordt, is Croonwolter&dros een aanjager van nieuwe technologische ontwikkelingen. Croonwolter&dros heeft de kennis, de ervaring en het vakmanschap om intelligente systemen en integrale, duurzame oplossingen te bedenken, te ontwikkelen, te implementeren, te beheren en te onderhouden.

Vanuit het motto 'Intelligentie door Technologie' draagt Croonwolter&dros met intelligente technologische elektrotechnische-, werktuigbouwkundige- en automatiserings- en informatisering oplossingen bij aan het duurzaam presteren van haar klanten. Met deze intelligente, zelfdenkende en -lerende systemen worden processen in of van objecten zoals schepen, tunnels, gebouwen en fabrieken efficiënter en effectiever.

Croonwolter&dros komt voort uit de fusie van Croon en Wolter & Dros. Croon was gespecialiseerd in elektrotechniek en automatisering. Wolter & Dros was gespecialiseerd in werktuigbouwkunde en meet- en regeltechniek. Met een rijke historie van 140 jaar Croon en 141 jaar Wolter & Dros zijn de ondernemingen samen goed voor meer dan 280 jaar ervaring in de techniek en opereren nu onder de merknaam Croonwolter&dros.

Met een omzet van 550 miljoen en 3.000 medewerkers is Croonwolter&dros het grootste in elektrotechniek, werktuigbouwkunde, automatisering en informatisering gespecialiseerde bedrijf van Nederland. Croonwolter&dros is actief in de volgende markten: Utiliteitsbouw, Industrie, Infra en Marine & Offshore en heeft vestigingen in heel Nederland.

Croonwolter&dros is onderdeel van de TBI holding.

De afdeling Croonwolter&dros STAF-Inkoop-Support is een subafdeling van de afdeling Inkoop. Het gehele inkoop afdeling bestaat uit ca. 38 medewerkers en de sub afdeling bestaat uit 10 medewerkers. De afdeling Croonwolter&dros STAF-Inkoop-Support geeft ondersteuning aan het inkoopproces en afdeling inkoop in breedste zin. De primaire processen binnen deze afdeling is o.a. zorgdragen voor:

- Analyses van het inkoopproces
- Het beschikbaar stellen van management rapportages.
- Datamanagement m.b.t. inkoopproces
- Procesoptimalisatie

De informatievoorzieningen die gebruikt worden binnen de afdeling Croonwolter&dros STAF-Inkoop-Support zijn datawarehousing, Power BI, Qlikview en een ERP (Enterprise Resource Planning) -pakket, Acto. De bedrijfscultuur in relatie tot de informatievoorziening is traditioneel/ conservatief.

2 De organisatie

In dit hoofdstuk worden de gegevens van de organisatie, de begeleiders en de school omschreven die betrekking hebben tot de afstudeerstage. Tenslotte wordt de organisatie en de afdeling inkoop & logistiek getoond in verschillende organigrammen.

2.1 Gegevens

Croonwolver&dros heeft verschillende vestigingen in het land. De opdracht wordt uitgevoerd door twee stagiaires. Om te verzekeren dat beiden afstudeerders hun B-competenties kunnen aantonen, is de opdracht verdeeld in twee verschillende onderdelen. Dit wordt verder uitgelegd in het hoofdstuk Mede-afstudeerder.

De afstudeerders werken op twee verschillende locaties, in Rotterdam en Amersfoort. Hier is voor gekozen omdat de medewerkers van de afdeling Staf-Inkoop-Support verspreid zitten over deze twee locaties. Beiden afstudeerders werken onder de Business Intelligence afdeling van Croonwolver&dros. De BI-afdeling implementeert de opdracht voor de afdeling Staf-Inkoop-Support.

Vestiging	Rotterdam	Amersfoort
Adres	Marten Meesweg 25	Amsterdamseweg 53
Postcode	3068 AV	3812 RP
Plaats	Rotterdam	Amersfoort
Telefoon	088 923 33 44	088 923 61 61

Tabel 27 Informatie Croonwolver&dros vestigingen

2.2 Begeleiders

De begeleider voor de afstudeerders binnen de organisatie is Rob Kabel. Rob is het aanspreekpunt wanneer er vragen, problemen en/of opmerkingen zijn. De eerste weken plant Rob (introductie) meetings in om de afdeling en de manier van werken binnen de organisatie te leren kennen. Tijdens de afstudeeropdracht blijft Rob betrokken bij de opdracht, om te controleren of de scope van de opdracht niet afwijkt.

De opdrachtgever van de afstudeeropdracht is Mathieu Koomans v.d. Dries. Mathieu geeft informatie over de opdracht en ondersteuning indien nodig. Aan het einde van de afstudeerperiode beoordeelt Mathieu de resultaten.

Achternaam opdrachtgever:	Koomans v.d. Dries
Voorletters opdrachtgever:	M.L. (Mathieu)
Titulatuur opdrachtgever:	Dhr.
Functie opdrachtgever:	Manager Procurement Support
Doorkiesnummer opdrachtgever:	+31 88 923 34 19
Email opdrachtgever:	mathieu.koomansv.d.dries@croonwolverendros.nl
Achternaam bedrijfsmentor:	Kabel
Voorletters bedrijfsmentor:	R.M.A. (Rob)
Titulatuur bedrijfsmentor:	Dhr.
Functie bedrijfsmentor:	Manager Business Intelligence
Mobiel bedrijfsmentor:	+31 6 12 91 03 94
Email bedrijfsmentor:	Rob.kabel@croonwolverendros.nl

Tabel 28 Croonwolver&dros begeleider en opdrachtgever

De docentbegeleider van de Hogeschool Leiden is Mischa Barthel. Mischa is het aanspreekpunt van school als er problemen en/of opmerkingen zijn. De begeleider komt minimaal één keer langs bij de organisatie. Dit (eerste) bezoek is in de tweede of derde week van de afstudeerperiode. Aan de hand van het bezoek in de tweede of derde week geeft de docentbegeleider een GO/NO-GO voor het afstudeerplan.

De examinerator van de Hogeschool Leiden is James Walls. De examinerator stelt het eindcijfer vast vanuit de Hogeschool Leiden. Dit wordt gedaan samen met de docentbegeleider en een externe beoordelaar op basis van de opgeleverde documenten, de verdediging en het advies van de bedrijfsbegeleider.

Achternaam examinerator:	Walls
Voorletters examinerator:	J.
Titulatuur examinerator:	Dhr.
Functie examinerator:	Docent informatica
Telefoonnummer examinerator:	+31611585667
Email examinerator:	Walls.j@hsleiden.nl
Achternaam docentbegeleider:	Barthel
Voorletters docentbegeleider:	M.
Titulatuur docentbegeleider:	Dhr.
Functie docentbegeleider:	Docent informatica
Mobiel docentbegeleider:	+31648133867
Email docentbegeleider:	barthel.m@hsleiden.nl

Tabel 29 Hogeschool Leiden begeleiders

De afstudeerder is verantwoordelijk voor het afstudeerproject en heeft zelf de opdracht geformuleerd en/of verzorgt. De afstudeerder is eigenaar van het afstudeerdossier. Tijdens het afstudeertraject kan de afstudeerder gebruik maken van de ondersteuning van Croonwolter&dros en Motion10 (Motion10, 2019). De afstudeerder is via Motion10 bij Croonwolter&dros terecht gekomen. Tijdens het afstudeertraject is Motion10 een informele ondersteuning voor de afstudeerders. Hier kunnen de studenten advies of ondersteuning vragen indien zij vastlopen in het project.

2.3 Mede-afstudeerder

Aan dit project werk ik samen met een andere afstudeerder. De opdracht is onderverdeeld in twee onderdelen. Mijn focus ligt bij de datakwaliteit en de data toegankelijkheid. Mijn verantwoordelijkheid is de data prepareren en controleren zodat het dashboard de werkelijkheid visualiseert. De andere component in het project is de visualisatie. Mijn collega-afstudeerder richt zich op de User Experience (UX) kant van het project. Dit omvat het ontwerpen van de rapporten en dashboards inclusief de functionaliteiten in Power BI.

Naam	S. van Dongen
School	Haagse Hogeschool
Opleiding	Business IT & Management

Tabel 30 Mede-afstudeerder informatie

2.4 Vorm van begeleiding

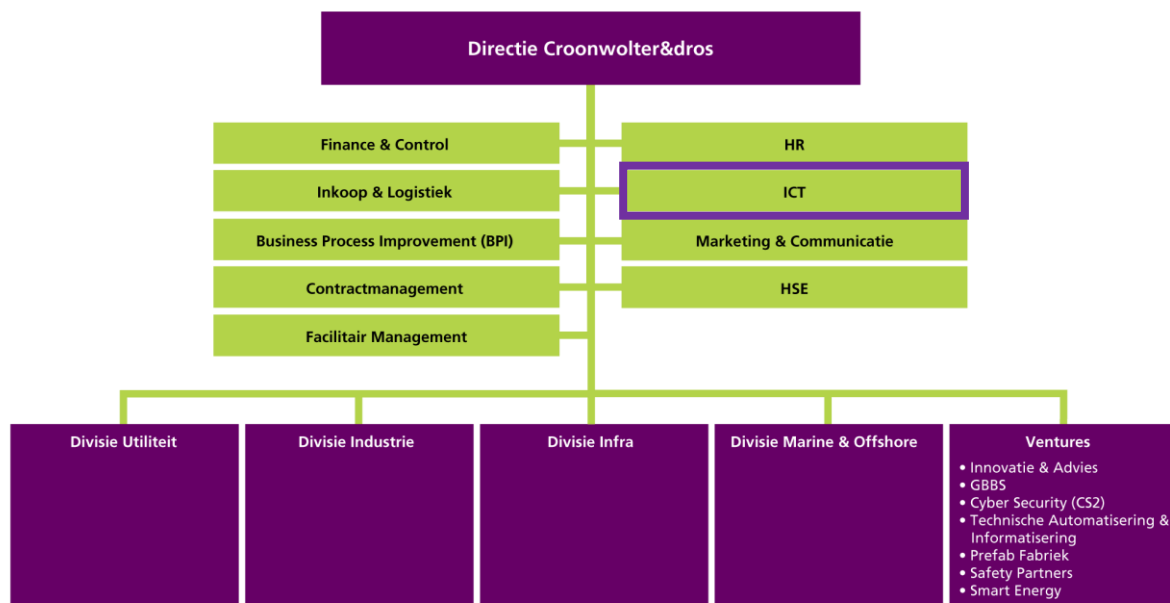
Rob Kabel is mijn bedrijfsbegeleider. Naast Rob helpen Mathieu (de opdrachtgever), Albert Huttenga (de datawarehouse-beheerder) of Susan Verbeij (Inkoper) met vragen/opmerkingen over de afstudeeropdracht.

Gedurende het afstudeertraject wordt er gewerkt volgens de SCRUM-AGILE methode. Er is afgesproken om twee keer per week een daily stand-up te hebben. Rob Kabel en beiden afstudeerders zijn altijd aanwezig tijdens de stand-ups. De opdrachtgever kan ook aanwezig zijn tijdens de daily stand-up, maar participeert niet actief mee.

Na elke sprint wordt er een sprint meeting gepland, waarbij het scrum-team en de Product-Owner aanwezig zijn. In deze meeting wordt er verteld wat er goed ging tijdens de sprint, maar ook tegen welke problemen het team is aangelopen en hoe die zijn opgelost. Het Scrum team laat de resultaten zien en beantwoordt hier vragen over. Tenslotte wordt er gekeken naar de belangrijkste items van het Product Backlog en wat er in de aankomende gerealiseerd wordt.

2.5 Organigram van de organisatie

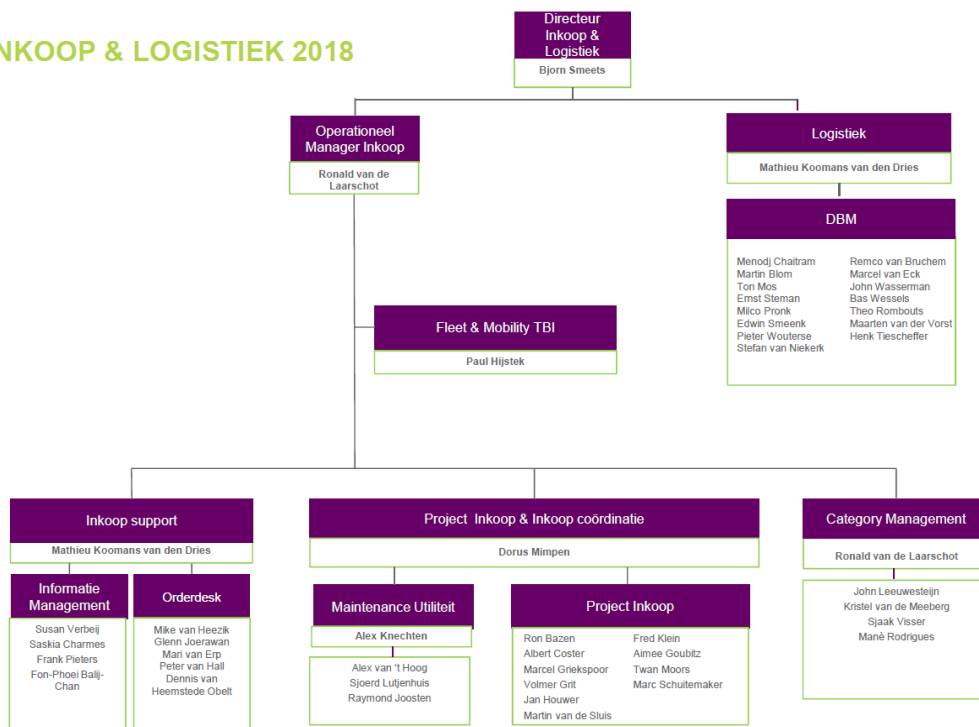
Onderstaand figuur toont de lijn-staf-organisatie structuur van Croonwolder&dros. De inkoopafdeling is deel van de staffuncties binnen de organisatie. Alle staffuncties ondersteunen de vijf onderliggende divisies. De divisie “Ventures” zijn kleine ondernemingen die ieder specifieke bijdragen leveren aan het functioneren van Croonwolder&dros.



Figuur 29 Organigram organisatiestructuur

Het organigram in figuur 2 toont de hiërarchische structuur van de inkoop- en logistiekafdeling.

INKOOP & LOGISTIEK 2018



Figuur 30 Organigram Inkoop & logistiek

2.6 Stakeholders

Bij het realiseren van het project zijn er stakeholders betrokken. Deze stakeholders zijn gedefinieerd in interne/externe stakeholders (Leansixsigma, Leansixsigma, 2019) en primaire/secundaire.

- De interne stakeholders staan dicht bij het project. Zij zijn direct betrokken en werken mee aan het eindresultaat.
- De externe stakeholders werken niet mee aan het project, maar zijn er wel bij betrokken en hebben belang bij het project.
- Primaire stakeholders zijn van redelijk tot groot belang voor het project.
- Secundaire stakeholders hebben geen direct belang bij het project, maar zijn wel van invloed op de implementatie van het project.
- De belangen van de stakeholders worden beschreven in de "Belang & Impact" kolom. De impact wordt aangegeven met een cijfer van 1 t/m 5.

Stakeholder	Intern/Extern	Primair	Functie	Belang & Impact
Mathieu Koomans v.d. Dries	Intern	Primair	Opdrachtgever	5
Susan Verbeij	Intern	Primair	Inkoper, consultant	5
Rob Kabel	Intern	Primair	Projectbegeleider	5
Jeroen Sman	Extern	Secundair	Informatie manager ICT	3
Managementteam	Extern	Primair	Stuurgroep	3
Albert Huttenga	Extern	Secundair	Informatieanalist	3
TBI Holding	Extern	Secundair	Overkoepelende onderneming van CW&D	2
BI-afdeling	Intern	Primair	Business Intelligence afdeling binnen CW&D	5
Sike v. Dongen	Intern	Primair	Afstudeerder	5

Tabel 31 stakeholders

3 Projectdoel

Dit hoofdstuk beschrijft wat er geïmplementeerd wordt met het afstudeerproject. De activiteiten die uitgevoerd worden om de producten te realiseren worden onderbouwt. De meerwaarde van het project voor de organisatie wordt in dit hoofdstuk duidelijk.

3.1 Bedrijfsdoelstelling

Het doel van dit project is om het huidige inkoopproces te optimaliseren. De output van dit project is om het dashboard van de inkoopafdeling van Croonwolter&dros, te modificeren naar een nieuw centraal dashboard met nieuwe functionaliteiten, hoge datakwaliteit en een real-time weergave. De huidige status van het dashboard is niet compleet en voldoet niet aan de eisen van de organisatie. Met het implementeren van deze oplossing wordt de inkoop overzichtelijker, doormiddel van het centraliseren en combineren van de data. Het combineren van de data maakt het mogelijk om nieuwe informatie en trends weer te geven, dit draagt bij aan het financieel- en performance beheer. Uiteindelijk wil Croonwolter&dros beslissingen nemen op basis van hoge datakwaliteit.

Croonwolter&dros werkt met Power BI, de inkoopafdeling werkt nog gedeeltelijk met Qlikview. Dit wil Croonwolter&dros omzetten naar Power BI. Door de huidige technologie om te bouwen naar de Power BI, wordt het mogelijk om de inkoopdata combineren met de rest van de organisatie.

Gedurende het traject zal de inkoopafdeling migreren naar Microsoft Azure. Het datawarehouse wordt vanuit een “eigen server” gemigreerd naar een Cloud-server van Microsoft Azure. Het Cloud-based platform biedt Croonwolter&dros nieuwe mogelijkheden om verder te ontwikkelen, met nieuwe API's en analysetools.

Het huidige systeem heeft bepaalde definities niet gestandaardiseerd. Hierdoor kunnen rapportages afwijken van de werkelijkheid. Met het standaardiseren van de definities binnen de organisatie ontstaat er een “single point of truth”. Dit voorkomt dat rapportages afwijken van de werkelijkheid.

4 Probleemstelling

Croonwolver&dros verwerkt meer dan 200.000 facturen per jaar en de inkoopafdeling behandelt 80% van de uitgaven van de organisatie. Inkoopdata is één van de basispilaren voor inzichtelijkheid en control van business performance en bewaking. Voor een goed bedrijfsbeleid en financieel beheer is het van belang dat de inkoop gecontroleerd en inzichtelijk is. Betrouwbare en tijdige data is van fundamenteel belang voor het beheren van processen en het behouden van overzicht over de verplichtingen. Dit is van toepassing op zowel operationeel als op directieniveau.

Naast het beschikbaar zijn van de data is het ook essentieel om te allen tijde te kunnen vaststellen hoe betrouwbaar de data is. De afdelingen hebben voor dezelfde situaties andere begrippen. Een leverancier bij de inkoopafdeling, is een crediteur binnen de financiën afdeling. Dit beïnvloedt de data integriteit. Het systeem binnen Croonwolver&dros heeft hier geen vast definities voor gedefinieerd.

Centrale dashboards die de inkoopafdeling ondersteunen bij het nemen van de juiste beslissingen en ook inzicht geven in de datakwaliteit zijn van invloed op de kostenbeheersing. Het huidige dashboard van de inkoopafdeling bevat geen real-time overzichten van de KPI's.

De inkoopafdeling werkt met bestelorderregels en de crediteuren afdeling werkt met factuurorders. De bestelorderregels bevatten de aantal producten/artikelen die besteld zijn bij één of meerdere leveranciers. De factuurorders bevatten de kosten van de kosten die gemaakt zijn door de inkoopafdeling. De twee datasets hebben een relatie met elkaar, elk factuur is berekend op de bestelorderregels. In het huidige datawarehouse worden de datasets niet gecombineerd met elkaar. Hierdoor kan Croonwolver&dros de omzet niet betrouwbaar berekenen en de huidige status van processen inzien.

Probleemstelling

Welke maatregelen dienen er gemaakt te worden om de data toegankelijkheid zodanig te verhogen, zodat de inkoopafdeling beslissingen kan nemen op basis van hoge kwaliteit data?

Deelvragen:

10. Welke KPI's en requirements zijn gedefinieerd binnen de inkoop- & logistiekafdeling?
11. Wie zijn de stakeholders en welk belang hebben zij bij het project?
12. Welke informatie wil de opdrachtgever aflezen van het dashboard?
13. Wat doet de organisatie aan data management?
14. Welke invloed heeft de data migratie naar Microsoft Azure op het project?
15. Welke aspecten van data management zijn belangrijk voor de organisatie?
16. Hoe wordt data management getoetst?
17. Op welke wijze kan de mate van de datakwaliteit inzichtelijk gemaakt worden in een dashboard?

5 Projectaannames

Gedurende het project worden er aannames gemaakt over verschillende aspecten waar de afstudeerders en de organisatie rekening mee dienen te houden. De aannames worden gespecificeerd in drie verschillende groepen:

- Uitsluitingen: Welke onderdelen binnen het project wordt niet aan gewerkt, of wordt er veronderstelt dat dit in orde is
- Afhankelijkheden: Waar is de organisatie van afhankelijk bij het realiseren van het project
- Middelen: Welke middelen zijn er nodig om het project uit te kunnen voeren.

5.1 Uitsluitingen

Ondanks dat de crediteuren data mart een onderdeel zijn van het project, ligt de focus van het project op de inkoopafdeling. De financiële afdeling heeft geen invloed op het project, enkel de data wordt gebruikt.

Datakwaliteit bevindingen worden aangegeven maar worden niet opgelost binnen het project. De afstudeerder geeft zijn bevindingen, advies en maatregelen die getroffen kunnen worden.

Andere datasets buiten Acto worden tijdens dit project niet gemoduleerd. Dit wordt wel intern gedaan door de organisatie, maar is geen onderdeel van dit project.

Het inkoopdashboard wordt gerealiseerd door een mede-afstudeerder. Hij doet onderzoek naar best-practices op het gebied van visualiseren van data en naar de requirements van de belanghebbenden.

5.2 Afhankelijkheden

Definities worden op verschillende niveaus gedefinieerd. Alle KPI's die voor de inkoopfunctie van belang zijn worden binnen dit project gedefinieerd, maar wellicht kunnen sommige KPI's niet binnen het project meegenomen worden, door data dat ontbreekt.

Indien de organisatie niet eenduidig het doel of de wensen kan aangeven, kan er geen oplossing geïmplementeerd worden waar de organisatie op kan bouwen.

5.3 Middelen

Om dit project te kunnen uitvoeren is het vereist om toegang te krijgen tot de gegevens die in de datawarehouse en data marts staan.

Om de wensen en eisen te achterhalen voor het project, zullen er verschillende workshops gehouden worden met de stakeholders. Op basis van de informatie uit de workshops zal er in overleg met de opdrachtgever en de projectbegeleider het project aangescherpt worden.

Tijdens het realiseren van het product wordt er gebruik gemaakt van verschillende soorten software. Het is vereist dat de projectleden toegang hebben tot deze software.

Software	OS	Omschrijving
Microsoft SQL Server	Elk apparaat	Databasebeheersysteem waar CW&D de data bewaard.
Microsoft Azure	Elk apparaat	Cloud computing-platform van Microsoft waar CW&D de datawarehouse host.
Power BI	Elk apparaat	Web-based visualisatietool van Microsoft. CW&D maakt met deze software rapportages en dashboards.
PowerPoint	Elk apparaat	Software om te presenteren.
Microsoft SharePoint	Elke apparaat	Platform van Microsoft voor het online uitwisselen van informatie en online samenwerken binnen de organisatie.
Azure DevOps	Elk apparaat	Microsoft online projectmanagement programma.

6 Deliverables

De producten en deelproducten worden bepaald aan de hand van de opdracht en de wensen van de opdrachtgever. De deelproducten maken deel uit van de school- en/of projectproducten.

ID	Titel	Deelproduct / Projectproduct / Schoolproduct	Keurder
D1	Afstudeervoorstel	Schoolproduct	Pitchcommissie, Klaas Jan Mollema
D2	Afstudeerplan	Schoolproduct	Mischa Barthel, James Walls
D3	Afstudeerverslag	Schoolproduct	James Walls, Mischa Barthel
D4	Adviesrapport	Projectproduct	James Walls, Mischa Barthel
D5	Implementatieplan	Projectproduct	Projectbegeleider
D6	Datakwaliteit dashboard	Projectproduct	Opdrachtgever, Projectbegeleider
D7	Externe data onderzoeksrapport	Deelproduct	Projectbegeleider
	Data kwaliteit rapport	Deelproduct	Opdrachtgever, Projectbegeleider
	Interviews	Deelproduct	Projectbegeleider
	Meetings	Deelproduct	Projectbegeleider
	Requirement analyse	Deelproduct	Opdrachtgever, Projectbegeleider
	Auditrapport	Deelproduct	Projectbegeleider
	Technisch ontwerp	Deelproduct	Projectbegeleider

Tabel 32 Deliverables

6.1 Productbeschrijvingen

Een nadere beschrijving van de deliverables inclusief de deelproducten. Indien een deliverable geen deelproducten heeft, is deze niet meegenomen in de productbeschrijving.

Deliverable	Beschrijving
D1 Afstudeervoorstel	Dit document is opgesteld naar aanleiding van de gesprekken met Croonwolter&dros. Dit product bevat geen deelproducten.

Tabel 33 Afstudeervoorstel

D2: Afstudeerplan	Beschrijving
D2.1 Analyse probleemstelling	De wensen en de eisen van de opdrachtgever achterhalen
D2.2 Stakeholdersanalyse	De stakeholders inclusief hun belangen achterhalen.
D2.3 Risicoanalyse	De risico's omschrijven inclusief de activiteiten om op een risico te reageren.
D2.4 Planning	De planning maken voor het project

Tabel 34 Afstudeerplan

D3: Afstudeerverslag	Beschrijving
D3.1 Bedrijfsprocessen analyseren	De bedrijfsprocessen in kaart brengen
D3.2 Technisch model	Een technisch model ontwerpen of toepassen aan het huidige technische model
D3.3 Datakwaliteit rapportage	De gebruikte methoden om de data kwaliteit te meten, inclusief een advies.
D3.4 Requirementanalyse	De requirements van de primaire stakeholders achterhalen.
D3.5 Interviews	Interviews afnemen van de medewerkers van de Inkoopafdeling.
D3.6 Meetings	Uitwerkingen van de meetings waar informatie is vergaard omtrent de opdracht.
D3.7 Auditrapport	De resultaten van de audits die uitgevoerd zijn.

Tabel 35 Afstudeerverslag

D4 Adviesrapport	Beschrijving
D4.1 Datakwaliteit Rapportage	Een analyse van de datakwaliteit inclusief een advies.

Tabel 36 Adviesrapport

D5 Implementatieplan	Beschrijving
D5.1 Requirementanalyse	De requirements van de primaire stakeholders achterhalen.
D5.2 Risicoanalyse	Een analyse van de risico's die zich voordoen bij het implementeren.
D5.3 Auditrapport	

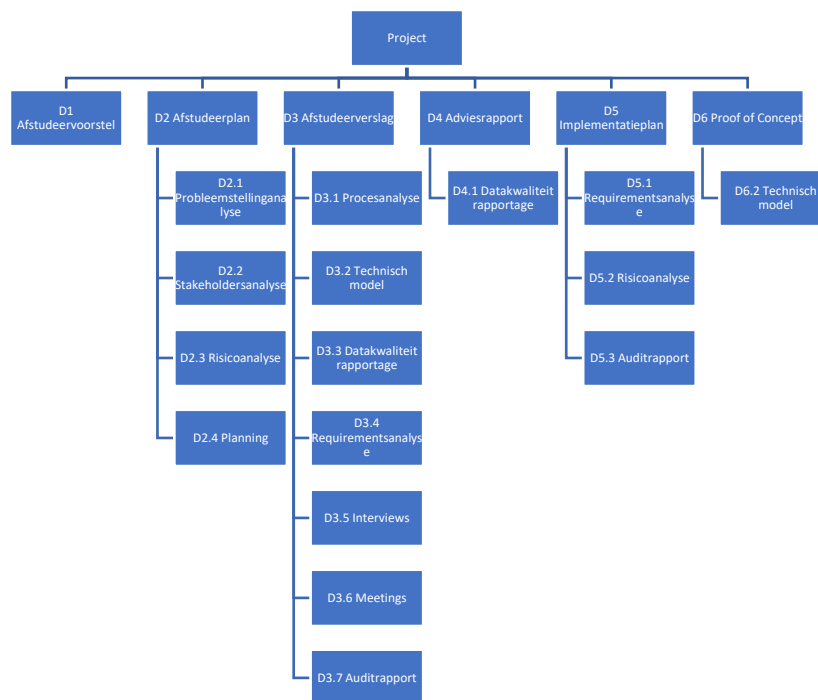
Tabel 37 Implementatieplan

D6 Datakwaliteit dashboard	Beschrijving
D6.2 Technisch model	Het model dat zich achter het dashboard bevindt.

Tabel 38 datakwaliteit dashboard

6.2 Product Breakdown Structure

De Product Breakdown Structure geeft een beter algemeen beeld van de opbouw van de producten. Het is een samenhang van alle deelproducten en deliverables.



Figuur 31 Product breakdown structure

7 Projectaanpak

Gedurende het project wordt er met de SCRUM-AGILE methode gewerkt. De SCRUM omgeving is in Microsoft DevOps ingericht. Tijdens het project zijn de studenten de SCRUM-Masters. De sprints zijn in iteraties weergegeven, met elke iteratie een andere focus. Iedere iteratie heeft een doorlooptijd van twee weken en na iedere sprint wordt er een sprint-review ingepland om de voortgangen te bespreken.

ITERATIES	ACTIVITEITEN
VOORONDERZOEK 12-02-2019 - 26-02-2019	Rechten, materiaal en software verkrijgen. Bedrijfsprocessen analyseren Stakeholderanalyse uitvoeren Probleem/kans definiëren Schrijven afstudeerplan
DEFINIËREN 26-02-2019 – 12-03-2019	KPI's achterhalen en definiëren Definiëren requirements Data onderzoeken Datakwaliteit meetinstrument inventariseren Literatuur onderzoek (best practices)
ITERATIE 1 13-03-2019 - 25-03-2019	Data management meetinstrument ontwerpen Data flow onderzoeken Data management methodieken onderzoeken Microsoft Azure verkennen
ITERATIE 2 26-03-2019 – 03-04-2019	Data management meetinstrument testen Microsoft Azure datakwaliteit services onderzoeken Eventuele nieuwe functionaliteiten onderzoeken Externe databronnen onderzoeken
ITERATIE 3 04-04-2019 – 19-04-2019	Afstudeerverslag opstellen Data management nulmeting uitvoeren Data management rapport opstellen Extra functionaliteiten realiseren Data stagen voor Power BI inkoopdashboard
ITERATIE 4 22-04-2019 – 30-04-2019	Acceptatietest Audits uitvoeren Testrapportage schrijven Inkoop dashboard presenteren aan opdrachtgever
NAZORG 01-05-2019 – 15-05-2019	Feedback inkoop dashboard verwerken Implementatieplan schrijven Afstudeerverslag afronden Presenteren resultaten datakwaliteit Verdediging afstudeerproject Datakwaliteit dashboard ontwikkelen

Tabel 39 Project activiteiten per iteratie

7.1 B-competenties

De B-competenties (hard skills) worden aangetoond doormiddel van de deliverables. Tijdens het ontwikkelen van de deliverables, wordt er rekening gehouden met de Rubric van de B-competenties van Business Data en Management.

B-competentie Niveau 3	Omschrijving
Bedrijfsprocessen analyseren	Om een passend meetinstrument te ontwerpen is een onderzoek vereist binnen de organisatie. Wat is belangrijk voor de organisatie en hoe gaat de organisatie te werk. Op basis van de benodigdheden van de organisatie wordt er onderzoek gedaan naar de best passende methoden, om het data management te meten. De methoden worden onderbouwt met literatuur onderzoek. De bevindingen van de meting(en) worden geformuleerd en een advies wordt gegeven met onderbouwing van de data. Indien een bedrijfsproces gewijzigd dient te worden, wordt hier een implementatieplan voor geschreven. De kwaliteit van de data wordt gevalideerd conform een methodiek voor data management. In het adviesrapport staan maatregelen die gemaakt dienen te worden om de datakwaliteit te waarborgen of te verbeteren.
Infrastructuur analyseren	De eerste weken wordt er onderzoek gedaan naar de huidige situaties met de (toekomstige) knelpunten waar de afdeling mee te maken heeft. Op basis daarvan worden de projectactiviteiten geformuleerd. Er wordt onderzoek gedaan naar best practices. De resultaten worden meegenomen in de onderbouwing van een oplossing. De mogelijke oplossingen worden in stappen beschreven hoe deze geïmplementeerd kunnen worden in de huidige infrastructuur. Met het combineren van verschillende data marts, moet er een gedeelte gemoduleerd worden. Hierbij hoort een technisch datamodel met de relaties tussen de tabellen. De nieuwe data infrastructuur moet de data kunnen combineren zodat het de compleetheid en de accuraatheid van de datakwaliteit verhoogt. De datakwaliteit wordt in een dashboard weergegeven, hierin is af te lezen hoe de huidige datakwaliteit is van de inkoopafdeling. Met deze informatie kan de afdeling actie ondernemen op basis van de data.
Bedrijfsprocessen ontwerpen	Voor het ontwerpen van het dashboard, worden er workshops gehouden om de wensen en eisen te achterhalen. De wensen en eisen worden samen met de opdrachtgever en met eigen uitgevoerde onderzoek uitgevoerd in een oplossing op maat voor Croonwolter&dros. Er wordt onderzoek gedaan naar de mogelijkheden om de datakwaliteit te testen. Dit kan in verschillende talen. Hierbij wordt goed rekening gehouden met de programmeertalen waar de organisatie kennis van heeft. Het implementatieplan bevat de stappen die genomen dienen te worden om het dashboard met functionaliteiten te integreren in de organisatie. Dit

	is inclusief de rollen die toegang hebben tot het dashboard. Het eindresultaat wordt gepresenteerd aan de opdrachtgever. Met het betrekken van de stakeholders bij het project, is het de bedoeling om een draagvlak te creëren, maar ook om de stakeholders die niet de waarde of het belang van het dashboard inzien, mee te krijgen.
--	---

Tabel 40 B-Competenties

7.2 A-Competenties

Tijdens het afstudeertraject moet de student aantonen dat hij de A-competenties (soft skills) beheerst om de opdracht adequaat op te zetten. De factoren van de A-competenties komen terug in het project.

A-competentie	Omschrijving
Onderzoek	Tijdens de eerste iteratie wordt er onderzoek gedaan naar de requirements en de wensen/eisen van de opdrachtgever. Om de requirements te achterhalen wordt de MoSCoW techniek gebruikt. De stakeholders en risico's worden volgens methoden verzameld.
Leren leren	De afstudeerder heeft weinig tot geen ervaring met datakwaliteit managen en meten. Om de kennis te ontwikkelen om het project succesvol te kunnen afronden is zelfstudie vereist. Het is mogelijk om ondersteuning te vragen, hiervoor moet de afstudeerder zelf initiatief nemen.
Professioneel werken	De KPI's worden doormiddel van een workshop achterhaalt. Deze methode wordt gebruikt om de belangrijkste, maar ook de minder belangrijke componenten van het project te definiëren. Op wekelijkse basis zijn er meetings met de opdrachtgever, stakeholders of met de begeleider. Er wordt gewerkt volgens een planning, waar de projectleden zich aan dienen te houden.
Innovatie	De afstudeerder dient een dashboard op te leveren die de huidige status van de datakwaliteit weergeeft. Op basis van de informatie op het dashboard moet er wel of geen actie ondernomen worden. Het dashboard wordt ontworpen door de afstudeerder en wordt gevalideerd door de opdrachtgever en bedrijfsbegeleider.

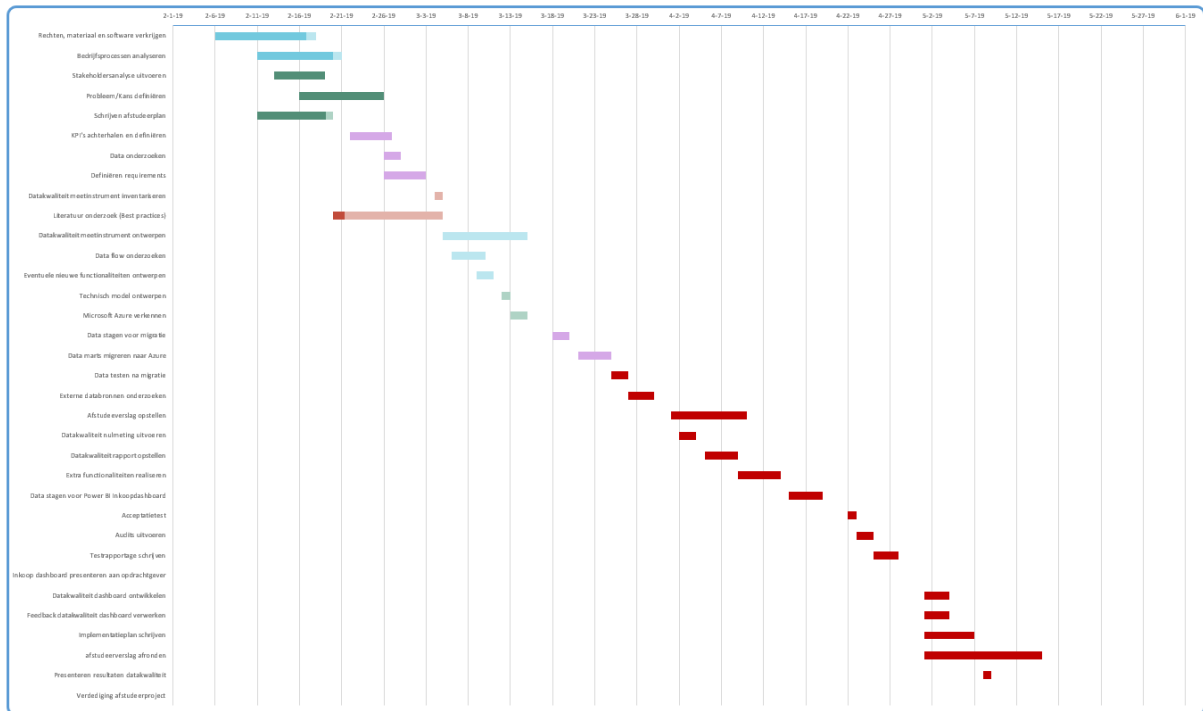
Tabel 41 A-Competenties

8 Projectplanning

Dit hoofdstuk weergeeft de planning in tabelvorm en in een gantt chart. Deze planning wordt wekelijks bijgehouden en als rode draad door het afstudeertraject gebruikt.

Iteratie	Task Name	Start Date	End Date	Duration (Days)	Days Complete	Days Remaining	Percent Complete
1	Rechten, materiaal en software verkrijgen	06-02-2019	18-02-2019	12	10,80	1,20	90%
1	Bedrijfsprocessen analyseren	11-02-2019	21-02-2019	10	9,00	1,00	90%
1	Stakeholdersanalyse uitvoeren	13-02-2019	19-02-2019	6	6,00	0,00	100%
1	Probleem/Kans definiëren	16-02-2019	26-02-2019	10	10,00	0,00	100%
1	Schrijven afstudeerplan	11-02-2019	20-02-2019	9	8,10	0,90	90%
2	KPI's achterhalen en definiëren	22-02-2019	27-02-2019	5	0,00	5,00	0%
2	Data onderzoeken	26-02-2019	28-02-2019	2	0,00	2,00	0%
2	Definiëren requirements	26-02-2019	03-03-2019	5	0,00	5,00	0%
2	Datakwaliteit meetinstrument inventariseren	04-03-2019	05-03-2019	1	0,00	1,00	0%
2	Literatuur onderzoek (Best practices)	20-02-2019	05-03-2019	13	1,30	11,70	10%
3	Datakwaliteit meetinstrument ontwerpen	05-03-2019	15-03-2019	10	0,00	10,00	0%
3	Data flow onderzoeken	06-03-2019	10-03-2019	4	0,00	4,00	0%
3	Eventuele nieuwe functionaliteiten ontwerpen	09-03-2019	11-03-2019	2	0,00	2,00	0%
3	Technisch model ontwerpen	12-03-2019	13-03-2019	1	0,00	1,00	0%
3	Microsoft Azure verkennen	13-03-2019	15-03-2019	2	0,00	2,00	0%
4	Data stagen voor migratie	18-03-2019	20-03-2019	2	0,00	2,00	0%
4	Data marts migreren naar Azure	21-03-2019	25-03-2019	4	0,00	4,00	0%
4	Data testen na migratie	25-03-2019	27-03-2019	2	0,00	2,00	
4	Externe databronnen onderzoeken	27-03-2019	30-03-2019	3	0,00	3,00	
5	Afstudeerverslag opstellen	01-04-2019	10-04-2019	9	0,00	9,00	
5	Datakwaliteit nulmeting uitvoeren	02-04-2019	04-04-2019	2	0,00	2,00	
5	Datakwaliteit rapport opstellen	05-04-2019	09-04-2019	4	0,00	4,00	
5	Extra functionaliteiten realiseren	09-04-2019	14-04-2019	5	0,00	5,00	
5	Data stagen voor Power BI Inkoopdashboard	15-04-2019	19-04-2019	4	0,00	4,00	
6	Acceptatietest	22-04-2019	23-04-2019	1	0,00	1,00	
6	Audits uitvoeren	23-04-2019	25-04-2019	2	0,00	2,00	
6	Testrapportage schrijven	25-04-2019	28-04-2019	3	0,00	3,00	
6	Inkoop dashboard presenteren aan opdrachtgever	30-04-2019	30-04-2019	0	0,00	0,00	
7	Datakwaliteit dashboard ontwikkelen	01-05-2019	04-05-2019	3	0,00	3,00	
7	Feedback datakwaliteit dashboard verwerken	01-05-2019	04-05-2019	3	0,00	3,00	
7	Implementatieplan schrijven	01-05-2019	07-05-2019	6	0,00	6,00	
7	afstudeerverslag afronden	01-05-2019	15-05-2019	14	0,00	14,00	
7	Presenteren resultaten datakwaliteit	08-05-2019	09-05-2019	1	0,00	1,00	
7	Verdediging afstudeerproject	31-05-2019	31-05-2019	0	0,00	0,00	

Figuur 32 Planning in Tabelvorm



Figuur 33 Planning in Gantt chart

9 Risicoanalyse

In overleg met de stakeholders is er een risicoanalyse uitgevoerd. Voor deze analyse is de methode van Fine & Kinney gebruik gemaakt. Met deze methode kun je doormiddel van berekende scores goed de risico's inschatten om te bepalen of er acties ondernomen moeten worden.

#	Risico	Bron	Risico beschrijving
R1	Te veel verschillende behoeften	Stakeholders	De verschillende stakeholders hebben hun eigen wensen en eisen voor het dashboard.
R2	Microsoft Azure performance	Microsoft	De Microsoft Azure performance voldoen niet aan de wensen en eisen van CW&D.
R3	Data marts zijn niet te combineren	Datawarehouse CW&D, Acto	De crediteuren data is niet te combineren met de inkoopdata. Hierdoor kunnen de wensen en eisen van de opdrachtgever niet worden geïmplementeerd.
R4	KPI's zijn niet aan te tonen met de beschikbare data	Stuurgroep, Datawarehouse CW&D, Acto	De KPI's die geformuleerd zijn, daarvan kan de status niet worden gevisualiseerd op het dashboard door gebrek aan data.
R5	Acto aanpassingen overschrijden budget	Acto	De aanpassingen aan de data marts kost te veel tijd, waardoor het te duur wordt voor CW&D. De kosten worden hoger dan de baten.

Tabel 42 Risico's binnen het project

Aanpak

De risico's en de omvang ervan wordt bepaald door een wiskundige formule. Om de risico-index te bepalen, worden er drie elementen meegerekend. De **kans** van het risico, de **blootstelling** van het risico en de **ernst**.

De formule is als volgt geformuleerd: **Risico = Kans x Blootstelling x Ernst** ($R = K \times B \times E$)

De kans-factor (**K**) staat voor de kans dat een risico zich voordoet. De kans geeft de verwachting weer en krijgt een referentieschaal van 0,5 tot 10.

De Blootstelling-factor (**B**) staat voor de mate waarin een risico opduikt. De Blootstelling krijgt een referentieschaal van 0,5 tot 10.

De Ernst-factor (**E**) is een aanduiding die de mogelijke schade en gevolgen weergeeft. De Ernst heeft een referentieschaal van 1 tot 100.

Waarden

De waarden van de factoren worden in onderstaande tabellen weergegeven.

Kans(K)	Waarde
Kan worden verwacht, bijna zeker	10
Goed mogelijk	6
Ongewoon, maar mogelijk	3
Alleen mogelijk op lange termijn	2
Zeer onwaarschijnlijk	1
Vrijwel onmogelijk	0,5

Tabel 43 Risico Kans

Bloodstelling(B)	Waarde
Voortdurend	10
Dagelijks tijdens werkuren	6
Wekelijks	3
Maandelijks	2
Enkele keren per jaar	1
Zeer zelden	0,5

Tabel 44 Risico Bloodstelling

Ernst(E)	Prioriteit
Catastrofaal	100
Ramp	40
Ernstig	15
Aanzienlijk	7
Belangrijk	1
Betekenisvol	0,5

Tabel 45 Risico Ernst

Afhankelijk van de berekende waarde voor Risico(R), moet er een conclusie getrokken worden voor het vervolg. De Risico index geeft weer hoe er gereageerd moet worden:

Risico cijfer	Urgentie	Omschrijving
$R < 20$	Laag	Aanvaardbaar
$R > 20 \text{ \& } R < 40$	Matig	Aandacht vereist
$R > 70 \text{ \& } R < 200$	Middel	Maatregelen vereist
$R > 200 \text{ \& } R < 400$	Hoog	Direct verbetering vereist
$R > 400$	Urgent	Werkzaamheden stoppen

Tabel 46 Risico index

9.1 Risicowaarden

Onderstaande tabel weergeeft de resultaten van de risico's. Het cijfer geeft de waarde van het risico weer. In de Risico index tabel op pagina 23 is af te lezen hoe zwaar de urgentie is per cijfer.

#	Risico	Cijfer (K*B*E)	Maatregel	Preventieve maatregel
R1	Veel verschillende behoeften	$3 * 6 * 7 = 126$	Samen met de opdrachtgever en projectbegeleider overleggen welke behoeften "must haves" en welke behoefte "nice to haves" zijn.	Eén persoon verantwoordelijk maken voor de requirements.
R2	Microsoft Azure performance	$3 * 2 * 15 = 90$	Contact opnemen met Acto en achterhalen hoe lang de servers offline zijn. Op basis van het tijdsbestek beslissingen nemen over het project.	De gewenste functionaliteiten testen in een Azure omgeving. Best practices onderzoek
R3	Data marts zijn niet te combineren	$1 * 2 * 40 = 80$	Dit is een ernstig risico, wanneer dit voorkomt moet er goed met de opdrachtgever overlegd worden hoe hij de data dan gemoduleerd wil hebben.	Innovatieve scripts ontwikkelen waardoor de factuur-bestelorderregels data bij de invoer aan elkaar gelinkt worden.
R4	KPI's zijn niet aan te tonen met de beschikbare data	$3 * 6 * 15 = 270$	In overleg met de opdrachtgever en de projectbegeleider wordt er gekeken naar de beschikbare data en de KPI's. Op basis daarvan wordt er een beslissing genomen over hoe het verder gaat.	De KPI's opstellen met de beschikbare data eraan te houden.
R5	Acto budget valt buiten de scope	$1 * 3 * 40 = 120$	Indien er over het afgesproken budget wordt gegaan, wordt er intern naar een oplossing gekeken. In het geval dat het niet intern opgelost kan worden, moet er een overleg plaatsvinden tussen de projectbegeleider en opdrachtgever.	-

Tabel 47 risicocijfer en maatregelen

10 Oplevering

De oplevering van het eindproduct is ingepland op 30 april 2019. Voor deze tijd moet het eindproduct getest en ge-audit zijn. De planning toont aan dat de het inkoopdashboard 30 april live gaat. Het inkoopdashboard wordt opgeleverd aan de Mathieu Koomans v.d. Dries. Mathieu representeert de inkoopafdeling. Doormiddel van een demo en presentatie worden de nieuwe functionaliteiten en visualisaties getoond.

Het datakwaliteit dashboard is het dashboard dat op operationeel niveau informatie visualiseert. Het datakwaliteit dashboard wordt ook ontwikkeld voor de inkoopafdeling, maar weergeeft andere informatie, voor andere gebruikers. Dit product wordt halverwege mei opgeleverd aan Mathieu Koomans v.d. Dries. Deze oplevering bevat ook de advisering rond data governance en de advisering van externe data.

De “schoolproducten” zoals dat in het hoofdstuk [Deliverables](#) staat beschreven worden naar de schoolbegeleider opgestuurd. Dit gebeurt in de week van 27-05-2019 tot 31-05-2019.

11 Bronvermelding

- Acto. (2019, 05 27). *Acto*. Opgehaald van Acto: <https://www.acto.nl>
- arbocatalogus, d. (2019, 02 14). *arbocatalogus*. Opgehaald van deArbocatalogus: <https://www.dearbocatalogus.nl/handreiking-intern-transport-en-logistiek/52-beoordelingstool-2-risicobeoordeling-van-fine-en-kinney>
- Brautigam, I. (2019, 02 19). *Scrumguide*. Opgehaald van Scrumguide: <https://scrumguide.nl/sprint-review/>
- DAMA. (2017). *DAMA DMBOK2*. -: DAMA.
- DataLadder. (2019, 05 25). *DataLadder*. Opgehaald van DataLadder: <https://www.Dataladder.com>
- ERP. (2019, 02 21). *ERPdefinitions*. Opgehaald van ERPdefinitions: <https://www.erpsystemen.nl/wat-is-erp>
- Fred Conijn, J. d. (2018, 01 01). *cmweb*. Opgehaald van cmweb: <https://cmweb.nl/2018/01/datamanagement-kwaliteit-gegevens-bepaalt-kwaliteit-informatie/>
- Jonker, R. (2019, 02 28). *Compact*. Opgehaald van Compact: <https://www.compact.nl/articles/datakwaliteitsonderzoek/>
- Kimball. (2019, 05 20). *KimballGRoup*. Opgehaald van KimballGroup: <https://www.kimballgroup.com/data-warehouse-business-intelligence-resources/kimball-techniques/dimensional-modeling-techniques/>
- Leansixsigma. (2019, 02 06). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/overzicht-stakeholders>
- Leansixsigma. (2019, 05 07). *Leansixsigma*. Opgehaald van Leansixsigma: <https://leansixsigmatools.nl/wat-zijn-key-performance-indactors>
- Microsoft. (2019, 4 20). *azure.microsoft*. Opgehaald van azure.microsoft: <https://azure.microsoft.com/nl-nl/overview/what-is-devops/>
- Motion10. (2019, 02 06). *Motion10*. Opgehaald van Motion10: <https://www.motion10.nl/>
- Oreilly. (2019, 03 14). *Oreilly*. Opgehaald van Oreilly: <https://www.oreilly.com/topics/oreilly-learning>
- Redactie. (2012, 10 26). *Business analytics*. Opgehaald van Business analytics: <https://www.business-analytics.biz/blogs/de-waarde-van-data-governance-en-het-zichtbaar-maken-ervan/>
- redactie, C. (2019, maart 22). *Computerworld*. Opgehaald van Computerworld: <https://computerworld.nl/business-intelligence/61874-wat-is-datawarehousing>
- Swaen, B. (2019, 04 20). *Scribbr*. Opgehaald van Scribbr: <https://www.scribbr.nl/onderzoeksmethoden/kwalitatief-vs-kwantitatief-onderzoek/>
- Tozzi, C. (2019, 02 28). *Syncsort*. Opgehaald van Syncsort: <https://blog.syncsort.com/2018/02/data-quality/how-to-measure-data-quality-7-metrics/>
- University, S. (2019, 05 09). *Stanford*. Opgehaald van Stanford: <https://web.stanford.edu/dept/pres-provost/cgi-bin/dg/wordpress/wp-content/uploads/2012/09/DG-Data-Definition-Best-Practices.docx>



- Verheul, M. (2019, 05 16). *Agilescrumgroup*. Opgehaald van Agilescrumgroup:
<https://agilescrumgroup.nl/welke-rollen-zijn-er-binnen-scrum/>
- Wethinknext. (2019, 02 07). *wethinknext*. Opgehaald van wethinknext:
<https://wethinknext.com/stakeholder-mapping/>
- Wikipedia. (2019, 03 07). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Databeveiliging>
- Wikipedia. (2019, maart 19). *Wikipedia*. Opgehaald van Wikipedia:
<https://nl.wikipedia.org/wiki/Data-architectuur>