

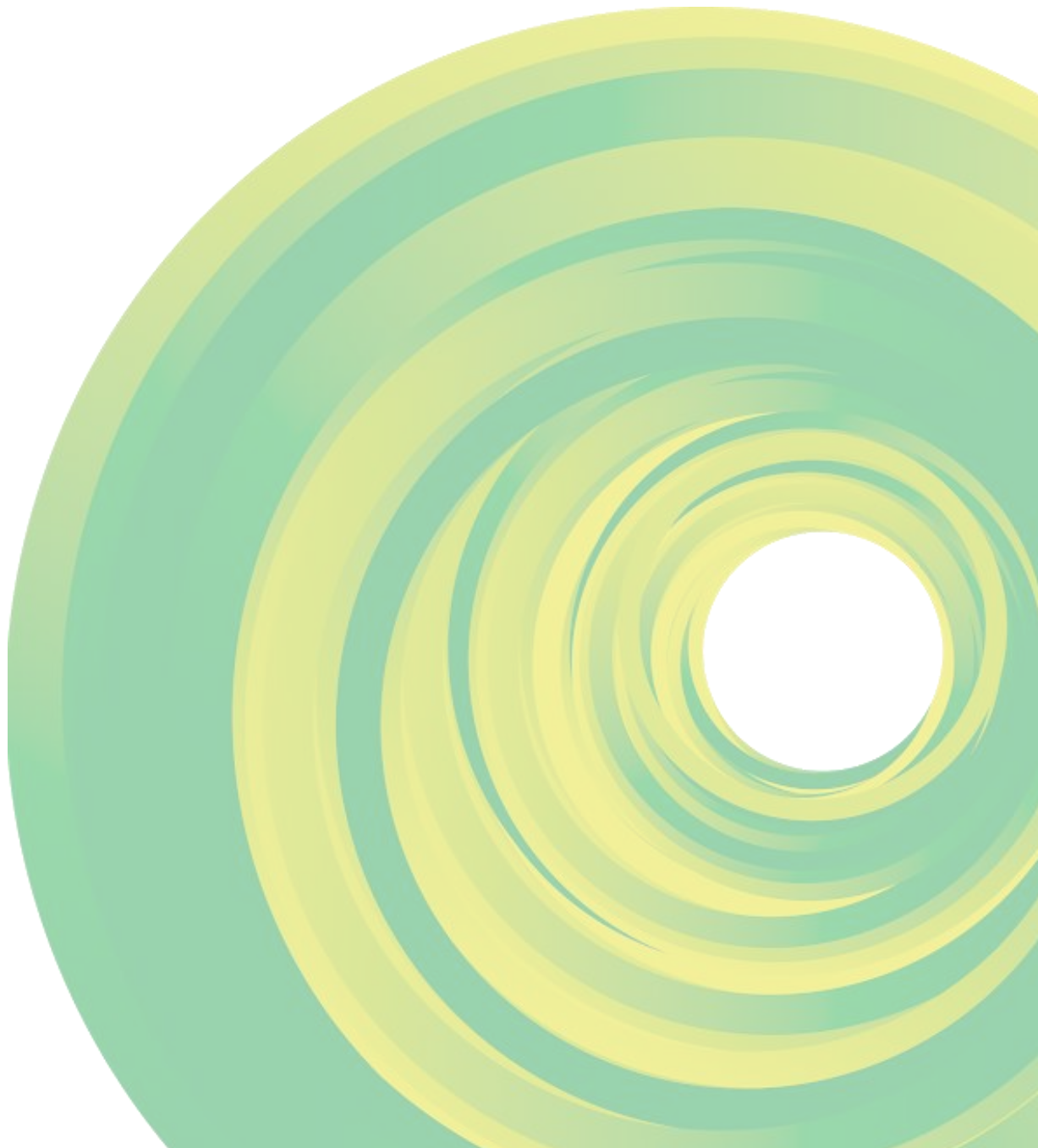
connect.
engage.
succeed.

SCRIPTIE:

TOEPASSING VAN DE
CORTANA INTELLIGENCE SUITE
VOOR DE INDUSTRIËLE SECTOR

**“Intelligence is not to
make no mistakes,
but quickly to see how
to make them good.”**

Bertolt Brecht



Versie:	1.0
Status:	Definitief
Datum:	3 juni 2016
Voorletters en naam:	J.M. Lelijveld
Telefoonnummer:	+316 122 122 72
Mailadres 1:	Marc.Lelijveld@Centric.eu
Mailadres 2:	Marc.Lelijveld@Me.com
Opleiding:	Business IT & Management
Opleidingsinstituut:	Haagse Hogeschool
Locatie:	Zoetermeer
Examinator 1:	M.J.M.R. Thissen
Examinator 2:	C.J.J. van Diest
Studieloopbaan begeleider:	P.A. Ritman
Start afstudeerperiode:	8 februari 2016
Eind afstudeerperiode:	3 juni 2016
Organisatie:	Centric Netherlands B.V.
Afdeling:	Advanced Solutions
Adres:	Rivium Westlaan 72
Postcode en plaats:	2900 AJ Capelle aan den IJssel
Opdrachtgever:	Erik Bouman
Functie opdrachtgever:	Line manager
Mailadres opdrachtgever:	Erik.Bouman@Centric.eu
Begeleider operationele zaken:	Narita Saija
Functie begeleider:	Consultant
Mailadres begeleider:	Narita.Saya@Centric.eu
Begeleider technische zaken:	Arjan Hordijk
Functie begeleider:	Software engineer
Mailadres begeleider:	Arjan.Hordijk@Centric.eu

SCRIPTIE:

TOEPASSING VAN DE CORTANA INTELLIGENCE SUITE VOOR DE INDUSTRIËLE SECTOR

Versiegeschiedenis:

Versie:	Wijzigingen:	Gewijzigd door:	Datum:
0.1	Eerste opzet t.b.v. tussentijdse feedback	J.M. Lelijveld	22 april 2016
0.2	Aanpassing n.a.v. tussentijdse feedback en volgende versie t.b.v. tussentijdse assessment	J.M. Lelijveld	12 mei 2016
0.3	Eerste volledige versie na verwerking feedback.	J.M. Lelijveld	20 mei 2016
0.4	Verwerking feedback Centric	J.M. Lelijveld	26 mei 2016
1.0	Eerste complete versie aangeboden voor eerste beoordeling,	J.M. Lelijveld	3 juni 2016

Tabel 1, Versiegeschiedenis

Copyright 2016, Centric Netherlands B.V.

Niets uit deze uitgave mag worden verveelvoudigd, opgeslagen in een geautomatiseerd gegevensbestand of openbaar gemaakt, in enige vorm of op enige wijze, hetzij elektronisch, mechanisch, door fotokopieën, opnamen of enige andere manier, zonder schriftelijke toestemming van Centric Netherlands B.V.

Referaat

Scriptie: Microsoft Cortana Intelligence Suite.
Lelijveld, J.M. (12016063)
Centric Netherlands B.V. – Juni 2016.

Dit document is opgesteld naar aanleiding van de door de student uitgevoerde afstudeeropdracht bij Centric Netherlands B.V. Dit afstudeerproject maakt onderdeel uit van de opleiding Business IT & Management aan de Haagse Hogeschool.

Gedurende de afstudeerperiode van zeventien weken is er onderzoek gedaan naar de Microsoft Cortana Intelligence Suite en de mogelijkheden die Centric hiermee kan bieden in de industriële sector. Hierbij is de organisatie Strukton Rolling Stock nauw betrokken voor een proof-of-concept voor de toepasbaarheid van de Microsoft Cortana Intelligence Suite. Het eindproduct van deze afstudeerstage is een proof-of-concept waarin de mogelijkheden van de Microsoft Cortana Intelligence Suite aangetoond worden.

Descriptoren

- Advanced Solutions
- Analytics
- Azure
- Azure IoT Suite
- BI
- Business Intelligence
- Centric
- Cortana
- Cortana Intelligence Suite
- Internet of Things
- IoT
- Machine Learning
- Microsoft
- Power BI
- Predictive Analytics
- Predictive Maintenance
- Rolling Stock
- Solution Engineering
- Strukton

Voorwoord

Voor u ligt de scriptie 'Toepassing van de Cortana Intelligence Suite voor de industriële sector.' Deze scriptie is geschreven in het kader van mijn afstudeeropdracht aan de opleiding Business IT & Management aan de Haagse Hogeschool te Zoetermeer.

Van 8 februari tot en met 3 juni 2016 ben ik bezig geweest met het onderzoek en het schrijven van mijn scriptie. Een onderzoek met daarbij een proof-of-concept wat de nodige tegenslagen heeft gekend op technisch gebied, maar toch tot een mooi eindresultaat heeft geleid.

Gedurende het uitvoeren van mijn onderzoek ben ik actief geweest binnen de afdeling Advanced Solutions bij Centric Netherlands B.V. Ik wil dan ook mijn opdrachtgever, Erik Bouman bedanken voor de kans die hij mij geboden heeft om als onderdeel van zijn team mijn afstudeeronderzoek te kunnen uitvoeren. Gedurende deze tijd heb ik veel geleerd en gehad aan de ervaringen van mijn begeleiders op zowel operationeel als technisch gebied. Daarvoor wil ik Arjan Hordijk, Cees Boonstoppel, Narita Saija en Rik van den Berg ook uitzonderlijk bedanken.

Uiteraard wil ik ook Han Geerligts, Marcel van der Linden en René van den Hoek van Strukton Rolling Stock bedanken voor hun medewerking aan mijn afstudeeronderzoek en de tijd die zij hebben geïnvesteerd in het verstrekken van domeinkennis en ondersteuning op technisch gebied.

De overige collega's bij Centric en medewerkers van Strukton Rolling Stock bedank ik voor de fijne samenwerking, ontvangen feedback en de mogelijkheden die jullie mij hebben geboden om te sparren.

Niet te vergeten wil ik mijn examinatoren, Maarten Thissen en Cees van Diest bedanken voor de opbouwende feedback en ondersteuning waar nodig.

Last but not least wil ik uiteraard ook mijn familie en vrienden bedanken die met mij mee hebben gedacht, verhalen aan hebben gehoord en ondersteund hebben.

Ik wens jullie veel leesplezier toe.

Marc Lelijveld,

Haastrecht, 3 juni 2016.

Inhoudsopgave

1	Inleiding	10
2	Het Bedrijf, Centric	11
2.1	CENTRIC	11
2.2	HET NIEUWE CENTRIC- MERK	11
2.3	KERNACTIVITEITEN	11
2.4	ADVANCED SOLUTIONS	12
3	De Opdracht	13
3.1	VERANDERING IN DE MAATSCHAPPIJ	13
3.2	SITUATIE	13
3.3	STRUKTON ROLLING STOCK	14
3.4	PROBLEEMDOMEIN	14
3.5	TE ONDERZOEKEN KWESTIE	16
3.6	HOOFDVRAAG	17
3.7	DEELVRAGEN	17
3.8	BEOOGD DOEL	17
3.9	DEFINITION OF DONE	17
4	Aanpak	18
4.1	WAAROM DEZE METHODE?	18
4.2	TOEPASSING VAN DE METHODE OP HET ONDERZOEK	18
5	Afbakening Industriële Sector	20
5.1	METHODE EN AANPAK	20
5.2	DEFINITIE INDUSTRIËLE SECTOR	21
5.3	NACE-CODE VAN STRUKTON ROLLING STOCK	21
5.4	AFBAKENING MET BETREKKING TOT HET ONDERZOEK	22
6	Bieb & Veld: Cortana Intelligence Suite	24
6.1	METHODE EN WERKWIJZE	24
6.2	CORTANA INTELLIGENCE SUITE	26
7	Veld: Belanghebbenden Bij Opdracht En Resultaat	29
7.1	METHODE EN WERKWIJZE	29
7.2	STAKEHOLDER MATRIX STRUKTON ROLLING STOCK	30
7.3	STAKEHOLDER MATRIX ALGEMEEN	33
7.4	CONCLUSIE	35
8	Veld: Betekenis Data	37
8.1	METHODE EN WERKWIJZE	37
8.2	DATA-ANALYSE	39
8.3	CONCLUSIE	42
9	Veld: Informatiebehoefte Strukton Rolling Stock	43
9.1	METHODE EN WERKWIJZE	43
9.2	VASTGESTELDE INFORMATIEBEHOEFTE	43
9.3	CONCLUSIE	45

<u>10</u>	<u>Werkplaats: Elementen Cortana Intelligence Suite</u>	<u>46</u>
10.1	METHODE EN WERKWIJZE	46
10.2	SOLUTION TEMPLATE	46
10.3	GEBRUIKTE ELEMENTEN UIT DE CORTANA INTELLIGENCE SUITE	48
10.4	CONCLUSIE	50
<u>11</u>	<u>Werkplaats: Visualisatie In Power Bi</u>	<u>51</u>
11.1	METHODE EN WERKWIJZE	51
11.2	VISUALISATIES IN POWER BI	52
11.3	POWER BI DASHBOARDS	52
11.4	CONCLUSIE	55
<u>12</u>	<u>Lab: Aansluiting Cortana Intelligence Suite Bij De Behoeften Van Strukton Rolling Stock</u>	<u>56</u>
12.1	WERKWIJZE	56
12.2	TOEGEVOEGDE WAARDE VAN BEREIKTE RESULTAAT	56
12.3	UITBREIDINGEN	57
12.4	CONCLUSIE	57
<u>13</u>	<u>Showroom: Aansluiting Cortana Intelligence Suite Voor De Industriële Sector</u>	<u>59</u>
13.1	TOEPASSING IN DE INDUSTRIËLE SECTOR	59
13.2	CONCLUSIE	60
<u>14</u>	<u>Conclusie En Aanbeveling: Cortana Intelligence Suite Voor Centric</u>	<u>61</u>
14.1	DOORONTWIKKELING	61
14.2	STANDAARDOPLOSSINGEN	61
14.3	AANBIEDINGSVORM EN INRICHTING	62
14.4	ONDERHOUD EN BEHEER	62
14.5	SAMENVATTING CONCLUSIE EN ADVIES	63
<u>15</u>	<u>Toekomstig Werk</u>	<u>64</u>
	<u>Literatuurlijst</u>	<u>65</u>
	<u>Verklarende Woordenlijst</u>	<u>68</u>
	<u>Lijst Van Figuren En Tabellen</u>	<u>71</u>
	<u>Bijlagen</u>	<u>72</u>

1 Inleiding

Als mensheid zijn wij beland in de volgende fase van het internettijdperk. Niet alleen computers en telefoons zijn onderling verbonden, maar alle dingen om ons heen gaan met elkaar communiceren, data verzamelen en autonoom acties uitvoeren. Dit wordt ook wel Internet of Things genoemd. De verzamelde data wordt steeds vaker gebruikt voor analyse achteraf, ook wel descriptive analytics. Wat is er gebeurd, en waarom is dat gebeurd?

Als we willen voorspellen wat er gaat gebeuren, spreken we predictive analytics. Op basis van historische data wordt patroonherkenning toegepast om hiermee gebeurtenissen in de toekomst te kunnen voorspellen. De Microsoft Cortana Intelligence Suite (voorheen Cortana Analytics Suite) bevat een set aan cloud diensten die deze vormen van analyse mogelijk maken.

Gedurende de periode van 8 februari tot 3 juni 2016 heb ik vanuit Centric Advanced Solutions een onderzoek uitgevoerd naar de mogelijkheden die de Cortana Intelligence Suite biedt. In het onderzoek is er gefocust op de industriële sector. Om de mogelijkheden van de Cortana Intelligence Suite in praktijk te kunnen testen, is Strukton Rolling Stock (hierna: Strukton) bereid gevonden om mee te werken aan dit onderzoek.

Binnen de scope van dit onderzoek is data van Strukton bij elkaar gebracht om hierop analyses uit te kunnen voeren en tot nieuwe inzichten te komen. Met deze proof-of-concept zijn de verschillende elementen uit de Cortana Intelligence Suite getoetst op toepasbaarheid. Voorspellingen op basis van data uit het verleden behoort ook tot de mogelijkheden van deze suite aan diensten. Echter is dit onderdeel in de loop van het project buiten scope geplaatst in verband met de beperkte beschikbare tijd.

In mijn scriptie wordt u stap voor stap meegenomen door het onderzoekstraject. Ieder hoofdstuk bevat een methode waarop dit specifieke onderdeel is aangepakt, gevolgd door de resultaten en tot slot een conclusie. Allereerst geeft hoofdstuk 2 nadere informatie over Centric als organisatie in de IT-branche. Aansluitend hierop wordt in hoofdstuk 3 de vooraf opgestelde opdracht behandeld. Vervolgens bevat hoofdstuk 4 de gekozen aanpak om tot dit eindresultaat te komen.

Hoofdstuk 5 geeft een afbakening aan van de industriële sector, waarbinnen de proof-of-concept is uitgevoerd. Waarna in hoofdstuk 6 informatie wordt gegeven over inhoud en mogelijkheden van de Cortana Intelligence Suite. Aangezien de proof-of-concept is uitgevoerd bij Strukton Rolling Stock, een externe partij, is het van belang om inzichtelijk te krijgen wie de stakeholders zijn bij de uitvoering van dit project. Dit wordt beschreven in hoofdstuk 7.

De inhoudelijke informatie met betrekking tot de proof-of-concept voor Strukton Rolling Stock wordt toegelicht in hoofdstuk 8 en 9. Beginnend met inhoudelijke kennis over de data van Strukton, gevolgd door de informatiebehoefte van de stakeholders. Hoofdstuk 10 beschrijft de gebruikte elementen uit de Cortana Intelligence Suite voor de proof-of-concept met in hoofdstuk 11 de resultaten hiervan en de wijze waarop deze bereikt zijn. De conclusie met betrekking tot de proof-of-concept wordt beschreven in hoofdstuk 12. Daarop aansluitend beschrijft hoofdstuk 13 de conclusie voor toepassing in de industriële sector.

Als afsluiting bevat hoofdstuk 14 conclusies en aanbevelingen voor Centric over de mogelijkheden en de toepassing van de Cortana Intelligence Suite.

2 Het bedrijf, Centric

In dit hoofdstuk wordt kort het afstudeerbedrijf toegelicht. Hierbij wordt informatie verstrekt over het ontstaan van Centric, het nieuwe Centric-merk en de kernactiviteiten van Centric. Verder wordt er ook kort informatie verstrekt over de afdeling waar de afstudeerstage is uitgevoerd.

2.1 Centric

Het onderzoek betreft een praktijkgerichte casus. Deze casus betreft een vraagstuk vanuit Centric, een van de grootste ICT- dienstverleners van Nederland maar ook actief in andere landen binnen Europa.

Sanderink Groep

Gerard Sanderink startte in 1992 de Sanderink Groep, de voorloper van het huidige Centric. Dit bedrijf groeide uit tot een van de grootste aanbieders van informatietechnologie in Nederland. Sinds januari 2000 is de naam Sanderink Groep veranderd naar Centric. Deze naam is afgeleid van het Griekse Kentrikos, wat centrum betekend. Centric groeide uit tot een volwaardige IT-leverancier op het gebied van detachering en software voor een aantal strategische markten, waaronder de lokale overheid, woningcorporaties, zorg, supply chain, bouw en de financiële sector. In deze sectoren biedt Centric continuïteit en totaaloplossingen. Tien jaar na de oprichting zette Centric de eerste stappen over de grens. Eerst in België, later ook in Duitsland, Zwitserland, Noorwegen, Zweden en Roemenië. Ook verbreedden de activiteiten zich en inmiddels biedt Centric naast IT-producten en –diensten, nu ook business process outsourcing en staffing services.

2.2 Het nieuwe Centric- merk

Sinds november 2013 worden alle Centric-activiteiten voortgezet onder de vlag van een nieuw Centric- merk. Een merk dat veel mogelijkheden biedt voor zowel klanten als medewerkers. Hiervoor worden professionals, partners en klanten verbonden, zodat zij hun talenten, kennis en ideeën met elkaar kunnen delen. Deze samenwerking resulteert in innovatieve, pragmatische oplossingen, die bijdragen aan verantwoorde groei van de organisatie en stabiliteit voor de klanten en leveranciers. Dit is de aanleiding geweest voor de nieuwe pay-off van Centric: 'Connect. Engage. Succeed.'



Figuur 1, Logo Centric

2.3 Kernactiviteiten

Centric speelde de laatste jaren actief in op de wens van organisaties om zich steeds meer te richten op hun eigen kernactiviteiten. Hoewel de oorsprong van Centric in de automatisering ligt, is Centric anno 2015 meer dan een IT-dienstverlener. De voorbije jaren is de dienstverlening verbreed en dat heeft in 2013 geleid tot een herinrichting van de organisatie met vier kernactiviteiten.

- Software Solutions

- IT Outsourcing
- Business Process Outsourcing
- Staffing Services

Business Process Outsourcing (BPO) is een relatief jonge activiteit voor Centric. Het aanbod groeit gestaag, van generieke bedrijfsprocessen zoals financiële of administratieve taken tot specifieke werkzaamheden voor financieel dienstverleners, pensioeninstellingen, gemeenten, woningcorporaties, en zorginstellingen. De BPO-klienten van Centric profiteren direct van de diepgaande kennis van branche specifieke bedrijfsprocessen, opgebouwd dankzij jarenlange ervaring met softwareoplossingen en staffing services.

Dankzij de bundeling in vier kernactiviteiten heeft Centric een overzichtelijk aanbod voor de markt en verbindt Centric kennis en expertise op effectieve wijze in één slagvaardige organisatie. Zo onderstreept Centric bovendien de ambitie om er op meerdere vlakken voor de klanten te zijn, waarbij de klanten kunnen rekenen op gesprekspartners die kennis hebben van hun bedrijfsprocessen en het aanbod van Centric kunnen vertalen in de specifieke klantsituatie. De klanten van Centric kunnen bij Centric terecht om hun IT-beheer uit te besteden, maar ook voor het inhuren van financiële professionals, voor complete branche specifieke softwareoplossingen of voor het uitbesteden van delen van de bedrijfsvoering. Centric heeft zichzelf hiermee in Nederland definitief als full serviceprovider gepositioneerd.

De focus ligt bij Centric op softwareontwikkeling, IT-outsourcing, business process outsourcing en staffing services. Binnen deze domeinen kunnen standaard of op maat gemaakte oplossingen afgenomen worden. Klanten vallen in diverse branches en sectoren, waaronder de bouw- en financiële sector, overheden, supply chain, onderwijs, waterschappen, woning coöperaties en de zorg.

(Centric Netherlands B.V., 2015)



Figuur 2, Kernactiviteiten Centric

2.4 Advanced Solutions

Centric Advanced Solutions houdt zich bezig met diverse op Microsoft technologie gebaseerde oplossingen waaronder de volgende oplossingen;

- SharePoint 2010/2013/SharePoint Online;
- Dynamics CRM 2011/2013/2015/2016/CRM Online.

(Centric Netherlands B.V. - Floris van Dijk, 2015)

Maar ook op Microsoft Azure gebaseerde oplossingen, maatwerk en ontwikkelingen in nieuwe technologieën behoort ook tot Advanced Solutions.

3 De opdracht

In dit hoofdstuk wordt de opdracht kort toegelicht zoals deze voorafgaand aan de afstudeerperiode is opgesteld. Allereerst wordt kort omschreven welke verandering er in de maatschappij plaatsvindt als aanleiding voor de opdracht. Deze maatschappelijke verandering wordt vervolgens nader toegelicht met een situatie omschrijving.

Als derde onderdeel van dit hoofdstuk wordt kort en bondig informatiegegeven over Strukton Rolling Stock, de betrokken organisatie bij deze afstudeeropdracht.

De te onderzoeken kwestie wordt in hoofdstuk 3.4 het probleemdomen kort toegelicht om vervolgens in hoofdstuk 3.6 tot een concrete onderzoeksvraag te komen. Deze onderzoeksvraag is in hoofdstuk 3.7 opgesplitst in diverse deelvragen.

Het beoogde doel wat met deze opdracht bereikt dient te worden, wordt beschreven in hoofdstuk 3.8, en aansluitend hierop de definition of done in hoofdstuk 3.9.

Het vooraf opgestelde afstudeerplan is terug te vinden in bijlage I met de daarbij behorende planning in bijlage II.

3.1 Verandering in de maatschappij

In de huidige maatschappij zijn we beland in het volgende industriële tijdperk. Niet alleen computers en telefoons zijn onderling verbonden, maar alle dingen om ons heen gaan met elkaar communiceren en autonoom acties uitvoeren. Alles wordt smart, zonder tussenkomst van de mens. Goed nieuws, of reden tot zorg voor Centric? Hoe dan ook: smart technologie gaat ons dagelijks leven ingrijpend veranderen. Van de manier waarop we wonen en betalen tot landbouw, transport en onze kleding. Hier ligt mogelijk een kans voor Centric om op deze verandering in te spelen.

3.2 Situatie

Op dit moment zijn veel apparaten en machines al uitgerust met diverse sensoren. Al deze sensoren leveren een grote hoeveelheid data op. Deze data geeft veel informatie over de status, functionaliteit en het proces van de machine. Met inzicht in deze data, kunnen apparaten en machines aangestuurd worden om bijvoorbeeld efficiënter ingezet te worden of defecten te voorkomen. Dit wordt Machine Analyses genoemd. Enkele termen op het gebied van smart industrie zijn:

- Data Science, interdisciplinaire processen en systemen om kennis op basis van data te verzamelen, zowel gestructureerd als ongestructureerd.
- Internet of Things, alledaagse voorwerpen verbonden met het internet.
- Big Data, een of meerdere databases van boven gemiddelde omvang, waarin 'big' een relatief begrip. Wat voor een kleine organisatie groot kan zijn, kan voor grote organisaties van gemiddelde omvang zijn.
- Content Analytics, toepassen van business intelligence op digitale content.
- Hybrid Cloud Computing, private en publieke hardware en software aangeboden via het internet.
- Cloud Computing, hardware en software aangeboden via het internet.

3.3 Strukton Rolling Stock

Om kennis op te doen van de Cortana Intelligence Suite, is Strukton Rolling Stock bereid gevonden om mee te werken aan het onderzoek. Strukton heeft een praktijksituatie waarbij verschillende elementen uit de Cortana Intelligence Suite op basis van de eigenschappen van deze suite een passende invulling kan bieden om te voldoen in de informatiebehoefte zoals deze in hoofdstuk 9 toegelicht wordt. De toepassing voor Strukton wordt als proof-of-concept gebruikt in dit onderzoek.

Achtergrond Strukton Rolling Stock

Strukton helpt klanten in de markten railsystemen, civiele infrastructuur en gebouwen verder door de toepassing van specialistische technieken. Behalve de Nederlandse en Europese markt bedient Strukton met zijn specialismen ook steeds meer de wereldmarkt. Het grotere Strukton bestaat uit meerdere werkmaatschappijen, waaronder Strukton Rail, die op zijn beurt weer onderverdeeld is. Strukton Rolling Stock vormt hier een onderdeel van.



Figuur 3, Logo Strukton Rail

De klanten van Strukton Rolling Stock, bevinden zich wereldwijd verspreid over de wereld. Zo heeft Strukton R.S.-klanten in onder andere India, Brazilië, Duitsland, Frankrijk, Italië en het Verenigd Koninkrijk. Strukton Rolling Stock heeft een unieke positie in de wereld als het gaat om de ontwikkeling, productie en levering van aandrijfsystemen (traction converters), boordnetomvormers (auxiliary converters) en voertuigbesturingssystemen (TCMS). Strukton zet zijn standaardmodules in voor klant specifieke oplossingen voor de renovatie van treinen, trams en metro's. (Werken bij Strukton, 2016) (Strukton Rail - Rolling Stock, 2016)

Om een beter beeld te krijgen de organisatie, is er in samenwerking met de medewerkers van Strukton Rolling Stock een Business Model Canvas ingevuld. Dit business model canvas is terug te vinden in bijlage III.

3.4 Probleemdomein

De probleemstelling kent twee kanten. Zo wordt de probleemstelling van Centric beschreven, en de probleemstelling van Strukton Rolling Stock, de organisatie die betrokken is bij het onderzoek voor de proof-of-concept.

Probleemdomein Centric

Binnen Centric Advanced Solutions is er geen ervaring met en kennis van machine analyses. Machine analyses bestaat onder andere uit bovengenoemde elementen op het gebied van smart industrie. Door de snelle ontwikkelingen in de maatschappij met daarbij een groeiende behoefte aan innovatieve oplossingen, is het van belang voor Centric Advanced Solutions om zich te ontwikkelen op dit gebied. Deze ontwikkeling is niet zo zeer noodzakelijk voor het voortbestaan, maar wel degelijk voor de positionering van Centric als IT-bedrijf. (Centric Netherlands, Solution Engineering, Will Loermans, 2015)

Smart industrie in de Gartner hype cycle

Enkele van de genoemde termen op het gebied van smart industrie zijn terug te zien in de Gartner hype cycle zoals deze getoond is in Figuur 4. De termen zoals deze terug te zien zijn in de hypecycle zijn de volgende:

- Affective Computing, studie en ontwikkeling van lerende en herkende systemen.
- IoT Platform, internet of things platform.
- Internet of Things, alledaagse voorwerpen verbonden met het internet.
- Machine Learning, kunstmatige intelligentie op basis van algoritmes.
- Hybrid Cloud Computing, hardware en software aangeboden via het internet.

Hierbij vertegenwoordigt het IoT Platform meerdere onderdelen van de eerdergenoemde termen op het gebied van smart industrie. Voor informatie over de Gartner hype cycle, zie bijlage IV.

Volgens het rapport van Gartner bevindt machine learning, predictive analytics en IoT zich in de Peak of Inflated Expectations fase, maar dicht tegen de Trough of Disillusionment fase aan. De huidige fase, Peak of Inflated Expectations, kenmerkt zich door veel bekendheid voor de technologie en een aantal succesverhalen. Echter staan hier wel een aantal mislukkingen en een lange doorlooptijd tegenover.

De aankomende fase, Trough of Disillusionment, kenmerkt zich door het geringe aantal succesverhalen ten opzichte van het aantal pogingen. Producenten van de technologie(en) steken veel tijd in de doorontwikkeling en gebruiken hierbij veel input van de early adopters. De early adopters betreffen de partijen die al gebruik maken van de technologie, al dan niet geslaagd.



Figuur 4, Gartner Hype cycle

(Gartner, 2015)

Kijkend naar de Gartner hype cycle is het bevorderlijk dat Centric zicht gaat ontwikkelen op het vlak van deze smart industrie. Er ligt een kans voor Centric om zich te ontwikkelen op onder andere het gebied van internet of things en machine learning om ervaring en kennis op te doen en de propositie aan te bieden aan haar klanten. Professor doctor Ben van Lier, tevens medewerker van Centric, is al jaren actief op het gebied van Internet of Things. Een uitbreiding hierop kan voor de organisatie van grote waarde zijn om meer bekendheid te genereren en de propositie te versterken. Tevens kan Centric zich als early adopter onderscheiden van de concurrenten door een oplossing te bieden gebaseerd op een technologie waar de concurrenten nog weinig tot geen kennis van hebben. Dit onderzoek dient ter verbreding van de kennis over machine learning waarbij de mogelijkheden voor de industriële sector onderzocht worden. (Centric - Prof. Dr. Ben van Lier, n.d.)

Voor Centric is de industriële sector relevant voor de toepassing van IoT om de oplossing te vermarkten. Om deze technologieën ook in praktijk te testen, is er een onderneming bereid gevonden om betrokken te worden bij het onderzoek voor een praktijk casus. Strukton Rolling Stock, gevestigd te Alblasterdam, is de betrokken partij. In hoofdstuk 3.3 Strukton Rolling Stock wordt de rol van Strukton in het onderzoek uitgebreid beschreven.

Probleemdomein Strukton Rolling Stock

De aandrijfsystemen, boordnetomvormers en voertuigbesturingssystemen van Strukton Rolling Stock beschikken over vele data sensoren. Deze data sensoren geven informatie over de status en de werking van deze onderdelen. Wanneer er bij Strukton of een van haar klanten een probleem optreedt met een van de onderdelen, kan de data informatie geven over de oorzaak van het probleem en mogelijke defecten.

In de huidige situatie wordt er bij een eventueel probleem een momentopname (snapshot) gemaakt van de data op het moment dat de fout zich voordoet. Hier wordt een x periode aan het begin en aan het eind toegevoegd om een mogelijke aanleiding te achterhalen, en een x periode aan het einde om het verloop van de fout te kunnen inzien. Deze periode voor en na de fout worden de slowsize en fastsize genoemd in de data. Daarbij wordt ook een slowfrequency en fastfrequency in de data genoemd. Deze waarden zijn benodigd om het aantal metingen in de periode voor en na de fout te kunnen berekenen.

Data-analyse

De registratie naar aanleiding van fouten wordt in de huidige situatie handmatig geanalyseerd door een medewerker van Strukton met de tool SCARP. Deze tooling is binnen Strukton ontwikkeld en in eigen beheer, enkel voor de analyse en van registraties. Op dit moment zijn de analyses en conclusies op basis van de interpretatie en ervaring van de medewerker. Onverhoopt kan het door de interpretatie en analyse van de medewerker voorkomen dat de gestelde analyse niet juist is. Hierdoor kan het zijn dat er onnodig of onjuist onderhoud gepleegd wordt en een structurele oplossing lang op zich laat wachten. De frequentie waarop een analyse fout verloopt is niet bekend.

Strukton ziet het beschreven proces graag geautomatiseerd waarbij de data snapshots automatisch geanalyseerd worden. Een welkome uitbreiding op de analyse is patroonherkenning waarmee voorspellingen van problemen en defecten gedaan kunnen worden. Deze inzichten helpen Strukton om hun klanten beter van dienst te zijn door continuïteit, beschikbaarheid en betrouwbaarheid van de geleverde producten en diensten te kunnen waarborgen. In de huidige situatie is dit nog een doel waarnaar gestreefd wordt door Strukton.

De probleemcasus en data van Strukton worden gebruikt voor een proof-of-concept met de Cortana Intelligence Suite als onderdeel van het onderzoek. Hiermee worden de theoretische eigenschappen van de Cortana Intelligence Suite in praktijk getest en biedt informatie voor Strukton om hun doel te kunnen bereiken.

3.5 Te onderzoeken kwestie

Centric, als Microsoft gold-partner, werkt veel met Microsoft gerelateerde producten en oplossingen. Zo wil Centric ook voor de ontwikkeling van machine learning, IoT en predictive analytics bij de technologieën van Microsoft blijven. Microsoft biedt hiervoor de Cortana Intelligence Suite.

Voor het onderzoek wordt gefocust op de industriële markt Deze markt is aangedragen door de initiatiefnemer binnen Centric, accountmanager Will Loermans. Om een praktijk casus te

kunnen onderzoeken, is Strukton Rolling Stock betrokken in het onderzoek. De probleemcasus van Strukton wordt hierbij uitgewerkt als proof-of-concept.

3.6 Hoofdvraag

Wat zijn de mogelijkheden van de Cortana Intelligence Suite voor Centric voor toepassing in de industriële markt?

3.7 Deelvragen

Voor de beantwoording van de hoofdvraag wordt antwoord gegeven op onderstaande deelvragen;

- Welk verschuiving zien we in de industriële sector waar Centric op in kan spelen?
- Wat zijn de mogelijkheden van de Cortana Intelligence Suite?
- Wat is het belang van Strukton van een geautomatiseerde data-analyse?
- Wie hebben er bij Strukton baat bij een geautomatiseerde data-analyse?
- Wat is de informatiebehoefte van Strukton waar een oplossing voor gezocht wordt in een geautomatiseerde analyse?
- Is de beschikbare data van Strukton geschikt om te kunnen voorzien in de informatiebehoefte?
- Sluit de Cortana Intelligence Suite aan bij de behoeften van Strukton Rolling Stock en de industriële markt?

3.8 Beoogd doel

Met dit onderzoek zijn meerdere doelen verbonden. Voor Centric is het van belang om de toepasbaarheid van de Cortana Intelligence Suite te toetsen. Hiermee wordt gekeken of Centric machine learning technologie aan kan bieden in de industriële markt.

Voor Strukton Rolling Stock is het aantonen van mogelijkheden tot geautomatiseerde data-analyse een doel waarnaar gestreefd wordt. Het probleemdomain van Strukton dient als proof-of-concept voor de toetsing van de toepasbaarheid van de Cortana Intelligence Suite. Voor Strukton wordt in de beperkte doorlooptijd van het onderzoek geen complete oplossing ontwikkeld. Een passende oplossing voor Strukton behoort tot een van de mogelijke vervolgstappen naar aanleiding van dit onderzoek.

3.9 Definition of done

In een periode van 17 weken een onderzoek doen naar de mogelijkheden van de Cortana Intelligence Suite. Om de mogelijkheden voor Centric in kaart te brengen wordt informatie gebruikt die door Microsoft naar buiten is gebracht, en direct in een proof-of-concept getest. Alle verzamelde en geconstateerde informatie wordt verwerkt in deze scriptie en de bijlagen.

Voor Strukton Rolling Stock wordt de proof-of-concept opgeleverd waarin de mogelijkheden van een geautomatiseerde data-analyse met de Cortana Intelligence Suite aangetoond worden.



5 Afbakening industriële sector

De onderzoeksvraag richt zich op de toepassing van de Cortana Intelligence Suite in de industriële sector. Hiervoor is het allereerst van belang om vast te stellen wat de industriële sector precies is. Hoofdstuk 5.1 beschrijft de gebruikte methode en aanpak waarop de industriële sector afgebakend is. Vervolgens wordt in hoofdstuk 5.2 de definitie van de industriële sector toegelicht. Aansluitend hierop worden de sub sectoren binnen de industriële sector (NACE) uitgewerkt. Aan het eind van dit hoofdstuk bij 5.3 vindt u de NACE-code van Strukton Rolling Stock. Tot slot wordt op basis van de NACE-codes een afbakening met betrekking tot het onderzoek toegelicht.

5.1 Methode en aanpak

In dit deelhoofdstuk wordt beschreven hoe tot de resultaten in dit hoofdstuk is gekomen, en welke methode en aanpak hiervoor gebruikt is.

Om het onderzoek af te bakenen, en scope creep te voorkomen, is dit hoofdstuk toegevoegd om duidelijkheid te geven waar het onderzoek zich op richt. De afbakening van de scope, en in dit geval het onderzoek, is een vorm van projectmanagement. Hierbij wordt vooraf de scope vastgesteld om te voorkomen dat er ongemerkt steeds (kleine) delen aan het onderzoek worden toegevoegd. Dit kan tot gevolg hebben dat de focus verloren gaat en er geen waardevol resultaat gerealiseerd kan worden. Dit kan veroorzaakt zijn doordat er is afgeweken van de werkelijk te onderzoeken casus.

Bij eventuele veranderingen in de scope is het van belang dat deze veranderingen in vastgelegd worden ter verantwoording. Ook is het van cruciaal belang dat de veranderingen worden besproken met de stakeholders. Dit heeft als doel dat de stakeholders op de hoogte blijven van de richting en het te verwachten eindresultaat.

Het is van belang dat de scope goed afgebakend is om de verwachtingen van de stakeholders te managen en tevreden te houden. Naar diverse bronnen is het voorkomen van scope creep in vijf stappen samen te vatten:

1. **Understand** what your client wants to achieve;
2. **Collaborate** with your client to produce a solution;
3. **Define** the scope of works that will provide that solution;
4. **Price** the solution in a manner that reduces the likelihood of overrun;
5. **Agree** the details of the service delivery in writing.

Aan het eind van dit hoofdstuk, bij Afbakening met betrekking tot het onderzoek, worden bovenstaande vijf punten toegelicht. Deze toelichting omvat de wijze waarop het gedurende dit onderzoek is toegepast.

(PMWIKI.nl, 2016) (Ewer, 2015) (Swart, 2010)

De afbakening van scope van dit onderzoek is voor aanvang voor het overgrote deel al gedaan. Dhr. W. Loermans, tevens initiator van het onderzoek, heeft in zijn documentatie over de propositie van machine analyses voor Centric de onderzoeksrichting bepaald. Hierbij is de industriële sector vastgesteld. Dit hoofdstuk is opgesteld om de scope vast te leggen en nader te omschrijven voor het onderzoek. Om meer informatie over de industriële sector te geven, wordt er in hoofdstuk 5.2 en 5.3 informatie verstrekt over wat er verstaan wordt onder de

industriële sector en hoe deze sector afgebakend is met betrekking tot de onderzoeksresultaten. Om deze afbakening inzichtelijk te maken, is gebruik gemaakt van de Europese standaard op het gebied van economische activiteiten die dient ter onderscheiding van sectoren.

(Centric Netherlands, Solution Engineering, Will Loermans, 2015)

5.2 Definitie industriële sector

De definitie van de markt Industrie volgens het CBS is als volgt:

Onder industrie wordt de mechanische, fysische of chemische verwerking van materialen, stoffen of onderdelen tot nieuwe producten verstaan. De verwerkte materialen, stoffen of onderdelen zijn grondstoffen uit de landbouw, bosbouw, visserij en mijnbouw, alsmede (half)fabricaten uit de industrie.

(Centraal Bureau voor de Statistiek, 2016)

De scheidslijn tussen industriële sector en andere sectoren is niet altijd even duidelijk. De industrie houdt zich bezig met de verwerking van materialen tot (nieuwe) producten. Ter verduidelijking volgt een aantal voorbeelden van activiteiten die in de sectie Industrie worden ingedeeld:

- Het pasteuriseren en bottelen van melk;
- Het drukken en aanverwante activiteiten;
- De productie van stortklare beton;
- De leerbewerking;
- De houtverduurzaming;
- Het galvaniseren, plateren, de warmtebehandeling van metalen en het polijsten;
- Het opnieuw bouwen of vervaardigen van machines (bijv. automotoren);
- Reparatie en renovatie van schepen;
- Loopvlakvernieuwing van banden.

Er zijn daarentegen ook activiteiten die weliswaar “nieuwe” producten opleveren maar niet tot de industrie worden gerekend. Daartoe behoren:

- Het optrekken van bouwwerken, de fabricage van constructies op de bouwplaats zelf (sectie F: Bouwnijverheid).

Activiteiten die niet leiden tot een nieuw product, maar tot een gewijzigde versie van hetzelfde product:

- Het opsplitsen van partijen stukgoederen in kleinere partijen, inclusief het verpakken, her verpakken;
- Het bottelen van producten als alcoholhoudende dranken of chemicaliën;
- De montage van computers volgens de specifieke wensen van een klant;
- Het sorteren van afval;
- Het mengen van verf.

Deze activiteiten worden ingedeeld in sectie G (groothandel en detailhandel).

(Centric Netherlands, Solution Engineering, Will Loermans, 2015)

5.3 NACE-code van Strukton Rolling Stock

Om economische activiteiten binnen sectoren op te delen wordt op Europees niveau een NACE-code gebruikt. NACE staat voor “Nomenclature statistique des Activités économique

dans la Communauté Européenne.” Ook wel: statistische naamgeving van de economische activiteiten in de Europese gemeenschap. In de gehele Europese Unie wordt de afkorting NACE als officieel woord gebruikt. De NACE-codes zijn bedoeld als hulpmiddel bij het opstellen van statistieken en overzichten. Voluit zijn de NACE-codes een geïntegreerd classificatiesysteem van economische activiteiten en producten.

Voor nadere informatie over NACE-codes, zie bijlage VI.

Bij het onderzoek wordt niet de gehele industriële sector betrokken. In het onderzoek wordt de organisatie Strukton Rolling Stock betrokken voor de ontwikkeling van een proof-of-concept. Met deze proof-of-concept worden de theoretische bevindingen bij de Cortana Intelligence Suite in praktijk getoetst. Meer informatie over Strukton Rolling Stock is terug te vinden in hoofdstuk 3.3.

Wanneer specifiek wordt gekeken naar de bedrijfsactiviteiten van Strukton Rolling Stock, is er raakvlak met meerdere NACE-departments. Dat zijn de volgende:

- C26.xx** Vervaardiging van computers en van elektronische en optische apparatuur.
Bij het ontwikkelen van de verschillende elektronische elementen (halffabricaten) ten behoeve van besturing en communicatie.
- C27.xx** Vervaardiging van elektrische apparatuur.
Bij de ontwikkeling van elektronische onderdelen ten behoeve van besturing en communicatie.
- C28.xx** Vervaardiging van overige machines en apparaten.
Het leveren van elementen (halffabricaten) voor spoorweg- en rolling stock producten, zoals treinen, trams en metro's.
- C30.xx** Vervaardiging van overige transportmiddelen.
Het leveren van elementen (halffabricaten) voor spoorweg- en rolling stock producten, zoals treinen, trams en metro's.

NACE-code Strukton Rail

In hoofdstuk 5.3 is aangegeven dat voor iedere onderneming maximaal één NACE-code toegekend kan worden. Omdat Strukton Rolling Stock onderdeel is van het grotere Strukton Rail, wordt de NACE-code van Strukton Rail gebruikt voor de gehele organisatie. Dit komt omdat voor de bepaling van de NACE-code de hoofdactiviteit van de organisatie wordt gebruikt.

Strukton Rail valt in een andere NACE-Section, en daarmee ook in een andere NACE-department en group. Strukton Rail heeft in de volgende NACE-code: F42.120.

- F - SECTIE F – Bouwnijverheid,**
 - 42 - Weg- en waterbouw,**
 - 421 - Bouw van wegen en spoorwegen,**
 - 4212 - Bouw van boven- en ondergrondse spoorwegen,**
 - 42120 - Bouw van boven- en ondergrondse spoorwegen.**

(Data.be - Business Info, 2016)

5.4 Afbakening met betrekking tot het onderzoek

Gedurende dit onderzoek wordt er gefocust op de industriële sector. Op basis van de NACE-codes valt Strukton Rolling Stock buiten de industriële sector. Dit komt omdat Strukton Rolling Stock onderdeel is van het grotere Strukton Rail. Zo heeft Strukton Rail meerdere

bedrijfsactiviteiten waarbij de hoofdactiviteit de NACE-code bepaald. Echter hebben de werkzaamheden van Strukton Rolling Stock voldoende raakvlakken met de industriële sector zoals beschreven in hoofdstuk 5.3.

Het onderzoek richt zich op de toepasbaarheid van de Cortana Intelligence Suite voor de gehele industriële markt. Dit omvat de gehele NACE-Section C – Industrie. Gedurende het onderzoek worden de resultaten in een proof-of-concept getoetst op basis van de behoeften van Strukton Rolling Stock. In hoofdstuk 13 wordt de aansluiting van de Cortana Intelligence Suite voor de industriële sector toegelicht. Daarbij worden ook andere toepassingsmogelijkheden toegelicht dan specifiek voor Strukton Rolling Stock.

Voorkomen scope creep

1. **Understand** what your client wants to achieve;
 - Om te begrijpen wat de betrokken partij, in dit geval Strukton Rolling Stock, precies wilt, worden er interviews gehouden met de betrokkenen binnen de organisatie. Ook wordt er extern georiënteerd op de mogelijkheden van de oplossingsrichting om met de betrokken partij mee te kunnen denken.
2. **Collaborate** with your client to produce a solution;
 - Om Strukton Rolling Stock continu op de hoogte te houden van de ontwikkelingen en voortgang, wordt er twee dagen per week op locatie bij Strukton gewerkt.
3. **Define** the scope of works that will provide that solution;
 - Om de scope te bewaken en voortgang te monitoren, worden er wekelijks voortgangsgesprekken gehouden met de direct betrokken personen bij Strukton. Bij deze overleggen is een duidelijke communicatie over haalbaar resultaat binnen de gestelde periode en het eventueel bijstellen hiervan iedere keer een van de gespreksonderwerpen.
4. **Price** the solution in a manner that reduces the likelihood of overrun;
 - De resultaten van het onderzoek worden gedeeld in de vorm van een onderzoeksrapport. De resultaten moeten helder omschreven worden waarbij ieder element toegelicht wordt. Hiermee wordt ernaar gestreefd om mogelijke vragen die kunnen ontstaan direct van een antwoord te voorzien.
 - De PoC wordt in een demo gepresenteerd aan de stakeholders waarbij de verschillende gebruikte elementen worden toegelicht en de werking wordt aangetoond.
5. **Agree** the details of the service delivery in writing.
 - De resultaten en genomen stappen om tot de oplossing te komen worden vastgelegd en besproken met de stakeholders om ervoor te zorgen dat de resultaten gewaarborgd zijn en reproduceerbaar zijn. Deze informatie is voor zowel Strukton als voor Centric van belang.



Figuur 6, Spoorweg

6 Bieb & veld: Cortana Intelligence Suite

Voorafgaand aan het onderzoek is in het onderzoeksplan vastgesteld dat de onderzoeksrichting de Microsoft Cortana Intelligence Suite bedraagt. Dit hoofdstuk geeft inzicht in wat de Microsoft Cortana Intelligence Suite exact is, wat de mogelijkheden zijn, en wat de richtingen zijn waarvoor het ingezet kan worden. Het onderzoeksplan is terug te vinden in bijlage VII.

In dit hoofdstuk wordt allereerst beschreven op welke wijze er onderzoek is gedaan, hoe ik te werk ben gegaan en welke methode hierbij toegepast is. Dit is beschreven in hoofdstuk 6.1 Methode en werkwijze. Daarop aansluitend beschrijft hoofdstuk 6.2 een nadere toelichting op de Cortana Intelligence Suite door toepassing van de aangegeven Methode en werkwijze.

6.1 Methode en werkwijze

Om informatie te verzamelen over de Cortana Intelligence Suite, is gebruik gemaakt van deskresearch en fieldresearch, ook wel veld en bieb in het framework van de methodekaart. De verschillende onderdelen uit het framework worden in hoofdstuk 4 Aanpak toegelicht. Om tijdens de deskresearch waardevolle informatie te vinden is er gebruik gemaakt van een lijst met termen waarnaar gezocht wordt. Onderstaand worden de verschillende zoekwoorden toegelicht en hoe hierop gekomen is.

Tevens is er informatie verkregen van een werknemer van Microsoft. In dit hoofdstuk wordt ook beschreven hoe de werknemer van Microsoft benaderd is, en waar over gesproken is. Ook is er voor de aanvang van de afstudeerperiode een event bijgewoond, waar door Microsoft, informatie met betrekking tot dit onderzoek verstrekt is.

Zoekwoorden deskresearch

Om een goed startpunt te maken met de deskresearch, is op basis van vooronderzoek een lijst met zoektermen samengesteld. Op basis van de eerste zoekresultaten is gebleken dat veel literatuur in het Engels geschreven is. Daarom heb ik ervoor gekozen om de lijst met zoektermen ook in het Engels op te stellen om met de resultaten van de zoekopdracht zo veel mogelijk nuttige informatie te vinden. Gedurende de deskresearch is de lijst met zoekwoorden uitgebreid naar aanleiding van gevonden resultaten.

Enkele gebruikte zoektermen zijn:

- | | |
|---------------------------------|------------------------------|
| • API | • Cortana |
| • Automated information systems | • Cortana Intelligence suite |
| • Automated systems | • Data mining |
| • Azure | • Data mining and ethics |
| • Azure analytics | • Decision tables |
| • Azure Data Catalog | • Decision trees |
| • Azure Data Factory | • Descriptive analytics |
| • Azure Data Lake | • Hadoop |
| • Azure Event Hub | • HDInsights |
| • Azure HDInsights | • Incremental learning |
| • Azure Machine Learning | • Instance based learning |
| • Azure SQL Datawarehouse | • Learning algorithms |
| • Azure Stream Analytics | • Learning-by-doing |

- Machine learning
- Machine learning algoritmes
- Machine learning API
- Mens versus machine
- Microsoft Cortana
- Microsoft Power BI
- Perceptual Intelligence
- Power BI
- Predictive analytics
- Prescriptive analytics
- Selffulfilling prophecy
- Supervised learning
- Unsupervised learning

De verzamelde informatie uit bovenstaande zoektermen is verwerkt in dit hoofdstuk en de hierop volgende hoofdstukken van dit document.

Om informatie te verzamelen, zijn er tijdens de deskresearch meerdere video's van zowel Microsoft als externe partijen over de Cortana Intelligence Suite bekeken. Dit zijn onder andere demonstratie video's en reviews van gebruikers. Deze video's hebben ook bijgedragen aan de kennis over de Cortana Intelligence Suite en de Azure IoT Suite.

Contact Microsoft

Informatie verkrijgen direct vanaf de bron is altijd het meest betrouwbaar en waardevol. Aangezien de Cortana Intelligence Suite een set aan producten van Microsoft is, en er binnen mijn persoonlijke netwerk contacten zijn met een medewerker van Microsoft, is er direct contact gezocht.

De aanleiding voor het contact met Microsoft was om antwoord te krijgen op enkele vragen die ontstaan zijn naar aanleiding van de deskresearch. Bij het contact zoeken met Microsoft, heb ik bewust gekozen voor een directe en persoonlijke benadering binnen mijn eigen netwerk. Zoals ik al meerdere malen ervaren heb bij voorgaande onderzoeken, reageren mensen niet snel en geven geen prioriteit aan de beantwoording van de gestelde vragen wanneer zij benaderd worden via social mediakanalen of via de mail. Ook is het vaak lastig om direct de juiste persoon te vinden en benaderen wanneer je deze persoon niet kent. Hiertoe is besloten om contact op te nemen met een medewerker van Microsoft binnen mijn eigen netwerk.

Dhr. R. Alders is benaderd met mijn vraag over de verschillen tussen de Cortana Intelligence Suite en de Azure IoT Suite. Dhr. Alders heeft mij vervolgens in contact gebracht met de juiste personen binnen Microsoft, dhr. B. van der Zwan, om de vragen te kunnen beantwoorden. De behoefte aan informatie over beide suites is ontstaan naar aanleiding de verwarring tijdens de deskresearch waarbij voor mij het verschil tussen de suites niet geheel duidelijk was.

Dhr. B. van der Zwan heeft telefonisch de tijd genomen om de vragen te beantwoorden en de verschillen tussen beide suites toe te lichten. Dit gesprek heeft in de vorm van een gestructureerd interview plaats gevonden. Daarbij waren de vragen vooraf al vastgesteld. Een interview een vorm van kwalitatief onderzoek waarbij gevraagd kan worden naar meningen en inzichten van de geïnterviewde. In het geval van dit interview waren er specifieke vragen waar dhr. B van der Zwan antwoord heeft kunnen geven op basis van de eigenschappen van de geboden diensten door Microsoft. Daarbuiten heeft dhr. B van der Zwan uit eigen inzicht nog een aantal aanvullingen gedaan op de gestelde vragen. Deze vorm van onderzoeken en informatie verkrijgen is fieldresearch, ook wel veld in de methodekaart. Het interview met dhr. B. van der Zwan is bijgevoegd in bijlage VIII.
(Right Markt onderzoek, sd)

Ook is er contact geweest met een collega die al eerder op het gebied van data analyse en machine learning voor Strukton Rail actief is geweest. Dit gesprek heeft een eerste aanzet gegeven voor de proof-of-concept. Een verslag van dit gesprek is bijgevoegd in bijlage IX.

Centric Tech Event

Het jaarlijks door Centric georganiseerde Tech Event, heeft plaats gevonden op 30 oktober 2015. Tijdens dit event worden de nieuwste ontwikkelingen van Microsoft gepresenteerd aan werknemers, relaties en klanten van Centric. Hierbij worden veelal gastsprekers van Microsoft betrokken om de informatie direct van de bron te verkrijgen.

Terugkijkend op deze dag is er ook veel informatie verstrekt met betrekking tot de Cortana Intelligence Suite. De presentatie van dhr. M. Westra van Microsoft is nogmaals bekeken en doorgenomen om hier relevante informatie voor dit onderzoek uit te halen.

6.2 Cortana Intelligence Suite

De Cortana Intelligence Suite, is een samenstelling van verschillende Clouddiensten van Microsoft. Al deze diensten worden aangeboden via de Cloud van Microsoft, deze wordt Azure genoemd. Binnen Azure zijn er vele verschillende mogelijkheden waar de elementen van de Cortana Intelligence Suite onderdeel van zijn.

Met de Cortana Intelligence Suite wordt het mogelijk gemaakt om gegevens om te zetten in intelligentie en actie. Het wordt mogelijk gemaakt om voorsprong op de concurrenten te behalen. Dit kan behaald worden door niet meer in de achteruitkijkspiegel te kijken en te analyseren wat er gebeurd is, maar te voorspellen wat er gaat gebeuren.



Figuur 7, Voordelen Cortana Intelligence Suite

(Microsoft - Cortana Analytics Suite, 2016)

Toepassing van de Cortana Intelligence Suite

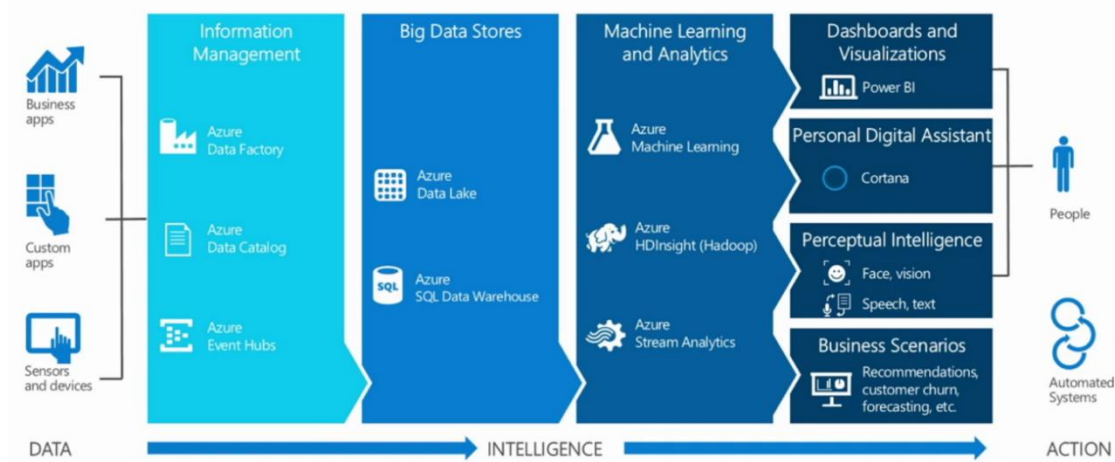
De Cortana Intelligence suite kan voor verschillende doeleinden ingezet worden in verschillende sectoren. Onderstaand een aantal voorbeelden.

- Productie-industrie, Voorspellen en preventief onderhoud mogelijk maken. Verbinding maken en proactief kapitaalgoederen kunnen bewaken en downtime elimineren.
- Financiële dienstverlening, Klanttevredenheid vergroten en online schalen. Effectiever verbindingen maken met gegevens en klanten.

- Detailhandel, Klanten trekken en zorgen dat ze terug blijven komen. Gepersonaliseerde service bieden.
- Gezondheidszorg, Positieve resultaten voor patiënten mogelijk maken. Verbinding maken met patiëntgegevens uit elke zorginteractie voor uitgebreide zorg.

Elementen van de Cortana Intelligence Suite

Zoals weergegeven in Figuur 8, is in de Cortana Intelligence Suite een deling gemaakt in drie grote groepen. Deze groepen kennen ieder meerdere onderdelen uit de Azure Stack, die weer meerdere mogelijkheden kennen.



Figuur 8, Elementen in de Cortana Intelligence Suite

Verklaring voor de elementen uit de Cortana Intelligence Suite worden beschreven in de verklarende woordenlijst en nader uitgewerkt in bijlage X.

Microsoft producten en technologieën vanuit Centric oogpunt

Centric Netherlands B.V. is sinds 2003 een gold certified partner van Microsoft. Dit is de hoogste rang partnerschap wat een organisatie met Microsoft aan kan gaan. Met dit partnerschap kan Centric snel inspelen op de nieuwe ontwikkelingen en veranderingen in de markt door continu gebruik te maken van de laatste nieuwe technologieën.
(Centric Netherlands B.V. - Hans Hengst - Partner Manager, sd)

Het partnerschap met Microsoft houdt ook in dat binnen Centric en de oplossingen die Centric verkoopt, voornamelijk Microsoft (gerelateerde) producten worden gebruikt en verkocht. Tevens is de afdeling Advanced Solutions, vanuit waar deze afstudeeropdracht is uitgevoerd, volledig Microsoft gericht. Afdeling Advanced Solutions wordt ook wel gezien als Microsoft Development afdeling.

Gartner's Magic Quadrant for Business Intelligence and Analytics Platforms

Gartner, een wereldwijd onderzoekende objectieve onderzoeksorganisatie binnen de IT, voert jaarlijks informatietechnologie onderzoeken uit om hiermee technologie gerelateerde inzichten te bieden aan haar klanten.
(Gartner, 2016)

Zo heeft Gartner een rapport uitgebracht waarin een rangschikking is gemaakt van de diverse aanbieders op het gebied van business intelligence en analytics platformen. Deze

rangschikking wordt getoond in het Magic Quadrant. Het Magic Quadrant is een matrix waar op de X-as de volwaardigheid van visie toont, en op de Y-as de mogelijkheid om de visie waar te maken. Om organisaties in te delen op de genoemde assen, wordt er onderscheid gemaakt in de volgende vier niveaus:

- **Leaders**, scoren het hoogste op beide criteria, zowel visie al de mogelijkheden om de visie waar te maken. Binnen deze categorie bevinden zich vaak grotere organisaties.
- **Challengers**, scoren hoog op de mogelijkheden om de visie te behalen, maar lager in de compleetheid van de visie.
- **Visionaries**, hebben een complete visie maar niet de capaciteiten om deze visie ook waar te maken.
- **Niche Players**, hebben geen complete visie en beschikken ook niet over de capaciteiten om deze visie waar te maken.

In februari 2016 heeft Gartner een rapport uitgebracht met daarin de meest recente inventarisatie en beoordeling van de organisaties die een business intelligence en analytics platform bieden. In dit rapport wordt met behulp van het Magic Quadrant getoond dat Microsoft op dit moment de leider is op het gebied van business intelligence en analytics platformen. Microsoft heeft hierbij de meest complete visie en capaciteit om deze visie ook waar te maken. Dit toont ook aan dat Microsoft op dit moment het meest vooruitstrevend is met business intelligence en analytics.

In het Gartner rapport wordt een verklaring gegeven voor de positie waar Microsoft in het kwadrant geplaatst is. Onderstaand citaat bevat de samenvatting van deze verklaring:



Figuur 9, Gartner's magic quadrant for business intelligence and analytics

Microsoft is the overall vision leader in the Magic Quadrant, having the greatest Completeness of Vision — as demonstrated by the commitment and focus of its BI and analytics leadership team to Power BI. Power BI is at the center of Microsoft's Azure story, which includes Azure Machine Learning, Azure HDInsight, Stream Analytics, and others, which solidifies its status as an area of continued strategic investment and innovation. Cortana Intelligence Suite is an umbrella for several of these offerings. Its market execution of Power BI 2.0 (released in July 2015) is greatly improved compared with its initial launch of Power BI, which failed to gain any traction in the market due to the confusing packaging, positioning and infrastructure prerequisites that few organizations could easily satisfy.

(Gartner - Josh Parenteau, Rita L. Sallam, Cindi Howson, Joao Tapadinhas, Kurt Schlegel, Thomas W. Oestreich, 2016)

Op basis van de conclusie van Gartner, kan geconcludeerd worden dat Microsoft business intelligence en analytics technologieën het goed doen en de beste positionering op de markt heeft. In combinatie met het gold-partnership voor Centric een reden om te ontwikkelen op dit gebied.

7 Veld: Belanghebbenden bij opdracht en resultaat

Met het resultaat van dit onderzoek worden meerdere belangen behartigd. Het belang van Strukton en het belang van Centric. In dit hoofdstuk worden de belanghebbenden van Strukton Rolling Stock nader toegelicht.

Allereerst wordt er in hoofdstuk 7.1 de Methode en werkwijze beschreven waarop tot de resultaten is gekomen. De resultaten van deze deelvraag wordt hierop volgend beschreven in hoofdstuk 7.2. Tot slot wordt er een Conclusie op deze deelvraag in hoofdstuk 7.4 beschreven.

7.1 Methode en werkwijze

Zoals toegelicht in hoofdstuk 3.3 is de aanleiding van het onderzoek naar de Cortana Intelligence Suite voor Strukton Rolling Stock het analyseren en voorspellen van problemen om deze in de toekomst te kunnen voorkomen. Het belang voor een organisatie lijkt hier dicht bij te liggen. Maar de belangen binnen de organisatie kunnen verschillend zijn. Dit hoofdstuk biedt meer inzicht in de belangen die spelen bij Strukton Rolling Stock voor een toepassing van de Cortana Intelligence Suite

Om deze informatie inzichtelijk te krijgen, zijn de directbetrokkenen bij de afstudeeropdracht geïnterviewd. Deze interviews hadden een kwalitatief karakter en ongestructureerd van aard. Hierdoor was het gemakkelijker om verklarende te stellen op basis van de gegeven antwoorden. Om geen belangrijke onderwerpen tijdens de interviews te vergeten, is er wel gebruik gemaakt van een korte lijst met onderwerpen die aan bod moeten komen gedurende het interview. Ook is iedere geïnterviewde gevraagd voor eventueel andere relevante personen om te interviewen.

(Alex Feijt - Research Company, 2013)

In het beginstadium van het onderzoek, werden in afwachting van de resultaten en richting van het onderzoek, nog geen organisatie brede bekendheid aan het onderzoek gegeven bij Strukton Rolling Stock. Hierbij is de keuze gemaakt om terughoudend te acteren voor het houden van interviews. Dit heeft geleid tot een beperkt aantal interviews met de directbetrokkenen. De verslagen van deze interviews zijn terug te vinden in bijlage XI en XII.

In later stadium van het onderzoek zijn concrete resultaten gedemonstreerd en het einddoel duidelijk gesteld voor de opdrachtgever binnen Strukton Rolling Stock. Op dit moment is ook gezamenlijk het besluit genomen om organisatie brede bekendheid te geven aan het onderzoek en de proof-of-concept die hierbij uitgevoerd wordt. Dit is ook het moment waarop feedback is gevraagd aan personen met verschillende rollen binnen de organisatie. Deze feedback was voornamelijk gericht op de aansluiting van de visualisaties bij de informatiebehoefte die ontstaat bij de betreffende functies.

Methode keuze

Om de stakeholders in kaart te brengen, is er een afweging gemaakt tussen verschillende matrixen en diagrammen. Er is een afweging gemaakt tussen de stakeholder matrix en de methode van Twynstra Gudde. (Werf &, sd) De theorie bijbehorend bij de stakeholder matrix wordt toegelicht in bijlage XIII.

Er is bewust niet voor de methode van Twynstra Gudde gekozen omdat deze niets zegt over de input van de stakeholders. De methode van Twynstra Gudde geeft inzicht in de betrouwbaarheid en belangen van de stakeholders, wat in mijn ogen dicht bij elkaar ligt. De invloed die uitgeoefend kan worden komt hierbij niet aan bod. De invloed van stakeholders kan cruciaal zijn voor het slagen of falen van het project of de opdracht, dit is naar mijn mening een belangrijk aspect wat niet overgeslagen mag worden.

De gebruikte methode is bekender en naar mijn mening gemakkelijker interpreteerbaar voor mensen die de theorie er achter niet kennen. Daarbij komt dat de matrix behorend bij de gebruikte theorie gemakkelijk hanteerbaar is tijdens het project om te beoordelen welke stakeholders wanneer en met welke frequentie op de hoogte gebracht moeten worden.

Werkwijze

Om de belanghebbenden in kaart te brengen, is er gedurende de interviews met de directbetrokkenen gevraagd naar de verschillende belangen die met de oplossingsrichting gemoeid zijn. De input van de respondenten heeft bij enkele gevallen geleid tot nieuwe interviews met de mogelijk belanghebbenden personen.

De betrokken stakeholders en hun belangen zijn vervolgens in een stakeholder matrix geplaatst. Er is bewust voor gekozen om niet de gehele theorie behorend bij de stakeholder analyse te volgen. Het analyse deel zoals dit in de theorie beschreven wordt, betreft het identificeren van de stakeholders. In mijn geval heb ik dat gedaan door de directbetrokkenen in interviews te vragen naar hun belang en mogelijke andere belangen die er spelen binnen de organisatie. Deze informatie heeft input gevormd voor de stakeholder matrix. Het doel wat ik wil bereiken met de stakeholder matrix is een overzichtelijke weergave en prioritering van de verschillende belangen. De volledige theorie behorend bij de stakeholder analyse is hierbij niet geheel gevolgd.

7.2 Stakeholder Matrix Strukton Rolling Stock

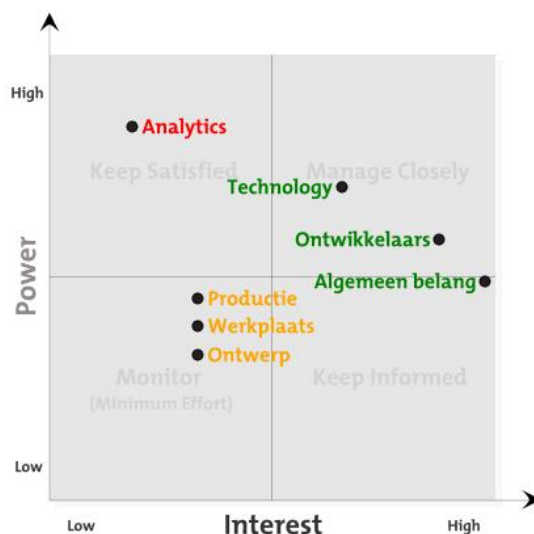
Zoals in voorgaand hoofdstuk toegelicht, wordt gebruik gemaakt van de stakeholder matrix. Deze matrix wordt gebruikt om de belanghebbenden van het onderzoek bij Strukton in kaart te brengen. Onderstaande stakeholder matrix tonen zowel functienamen als personen. Deze keuze is gemaakt omdat niet van iedere persoon en functie een naam bekend is of meerdere personen vertegenwoordigd.

In de stakeholder matrix is er gebruik gemaakt van onderstaande kleurcoderingen:

- **Groen:** Supporters en steun gevende.
- **Geel:** Neutraal.
- **Rood:** Critici en tegenstanders.

Stakeholder matrix Strukton Rolling Stock

Onderstaande stakeholder matrix toont een prioritering van de stakeholders op functieniveau. Hierbij worden enkel functies gebruikt van mogelijk belanghebbenden.



Figuur 10, Stakeholdermatrix Strukton Rolling Stock

Technology

Medewerkers bij Strukton Rolling Stock die zich bezighouden met de technologie, tonen gedurende het project belangstelling. Deze medewerkers denken mee gedurende de ontwikkeling, en zijn bereid om hier tijd in te steken. Vanuit zowel persoonlijke als business gedreven interesse houden zij de ontwikkelingen bij. Enkele motiverende factoren voor deze medewerkers zijn de mogelijkheden tot doorontwikkeling en toekomstige toepassingen.

Ontwikkelaars

Een van de werkzaamheden van Strukton Rolling Stock is het ontwikkelen van software voor besturing van locomotieven. Eventuele fouten die gedurende het gebruik van de locomotief optreden en in de registratie data terug te vinden zijn, kunnen veroorzaakt worden door onjuiste aansturing. Vanuit het oogpunt van de ontwikkelaars is het interessant om te weten waar de software fouten vertoont en hoe vaak deze fouten voorkomen. Op basis van deze informatie kan de software, specifiek gericht op de fouten, doorontwikkeld worden.

De ontwikkelaars bij Strukton Rolling Stock zien hiermee een meerwaarde in de toepassing van de Cortana Intelligence Suite op de registratie data. Hierbij tonen de ontwikkelaars interesse in de voortgang en de mogelijkheden. Hierdoor is de interesse van de ontwikkelaars hoog. Ook hebben de ontwikkelaars invloed op het eindresultaat door hun informatiebehoefte kenbaar te maken. Deze informatiebehoefte wordt meegenomen in de toepassing van de Cortana Intelligence Suite.

Algemeen belang

Het algemeen belang bij Strukton Rolling Stock, ook wel organisatorisch belang, speelt ook een rol. Strukton Rolling Stock beschikt over grote hoeveelheden data waar mogelijk meer informatie uit te halen is, dan op dit moment gedacht wordt. Deze informatie kan indirect de kwaliteit van producten en diensten, maar ook de positionering van Strukton op de markt verbeteren.

Onderwerpen als big data, remote monitoring, analytics en machine learning zijn onderwerpen waar hedendaagse organisaties druk mee bezig zijn en veel in het nieuws komen. Deze onderwerpen krijgen veel aandacht en kunnen futuristische toepassingen brengen. Wanneer men te laat begint met deze ontwikkelingen, kan dit een achterstand ten opzichte van de concurrenten betekenen en mogelijk zelfs verlies van marktaandeel. Voor Strukton Rolling Stock gronden om de ontwikkeling op hiervoor genoemde vlakken tijdig in gang te zetten.

Om bovengenoemde reden heeft Strukton Rolling Stock veel belang bij de PoC met de Cortana Intelligence Suite. De kennis en ervaring die hieruit voortvloeit, kunnen richtinggevend zijn voor bredere ontwikkeling binnen Strukton Rolling Stock. Hierdoor is de interesse van deze groep personen hoog. Zij zijn erg benieuwd naar de bereikte resultaten en de mogelijkheden. De invloed van deze groep heeft een mindere omvang. Er kan door de medewerkers invloed uitgeoefend worden door mee te denken over oplossingsrichtingen en andere toepassingen, maar hebben hierin geen doorslaggevende rol.

Productie

Productiemedewerkers bij Strukton hebben minder belang bij het analyseren van de remote monitoring data. Deze personen staan neutraal ten opzichte van het resultaat.

De interesse van productiemedewerkers ligt niet zo zeer bij het project en de vordering, maar meer bij het ontwerp en de werking van de remote monitoring box waarmee de data verzameld wordt. Vanuit de positie van een productiemedewerker is er nauwelijks tot geen invloed op het resultaat van de analyse werkzaamheden.

Het uiteindelijke resultaat van analyse werkzaamheden kunnen wel waardevol zijn voor productiemedewerkers. Zo kan de analyse mogelijk aantonen dat bepaalde fouten vaak voorkomen op een onderdeel uit een bepaalde serie. Voor productie kan dit aanleiding zijn om het probleem met dit onderdeel nader te analyseren en het product te verbeteren in samenwerking met de ontwerp medewerkers.

Ook kan de foutanalyse bijdragen aan het voorraadbeheer van de onderdelen. Wanneer een specifiek onderdeel geregeld defect gaat, kan dit aanleiding zijn om dit onderdeel meer op voorraad te nemen om de problemen sneller op te kunnen lossen.

Werkplaats

Werkplaatsmedewerkers hebben eveneens als de productiemedewerkers gering belang bij de uitkomsten van de analysewerkzaamheden met de remote monitoring data. Hun invloed is ook gering. Hierdoor is er voor deze doelgroep minder inspanning nodig om hen op de hoogte te houden van de vorderingen.

De uitkomsten van de analyses zijn voor de werkplaatsmedewerkers op langere termijn wel van belang. Zo is het denkbaar dat werkplaatsmedewerkers op basis van de uitkomsten van machine learning defecten aan zien komen. Op deze wijze kan onderhoud tijdig ingepland worden om defecten tijdens gebruik van de locomotief te voorkomen. Dit wordt ook wel predictive maintenance genoemd. Dit is pas op langere termijn toepasbaar aangezien de patroonherkenning om de fout aan te zien komen ingeleerd dient te worden. De machine learning aspecten in de oplossing dienen voorzien te worden van data. In deze data moet worden aangegeven wanneer een onderdeel defect is, voordat het patroon wat leidt tot een defect gelokaliseerd kan worden in de data.

Ontwerp

Medewerkers bij Strukton die de producten ontwerpen hebben een gelijk belang als de werkplaatsmedewerkers. In eerste instantie hebben de ontwerp medewerkers nauwelijks belang bij de uitkomsten van de analyse op korte termijn. Op langere termijn hebben ontwerpmedewerkers wel degelijk belang bij de analyses. Wanneer een specifiek onderdeel geregeld defecten vertoont, kan dit op langere termijn, afhankelijk van de oorzaak, leiden tot herontwerp om de fouten te voorkomen in toekomstige situaties.

Analytics

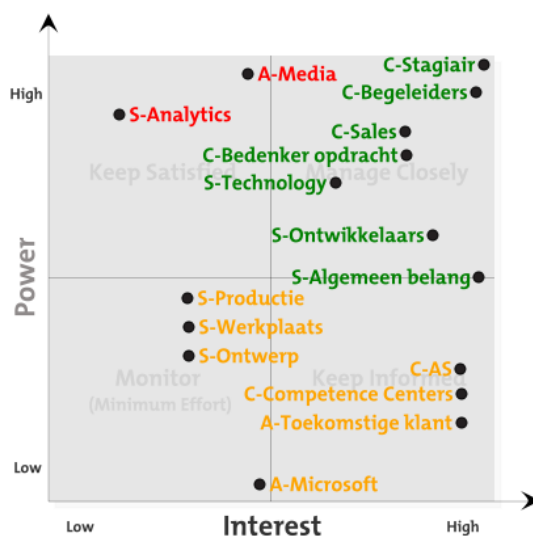
Binnen Strukton Rolling Stock worden de registraties in de huidige situatie handmatig geanalyseerd wanneer er een fout optreedt. De personen die deze analyses uitvoeren, kunnen de ontwikkeling en analyse van de registratie data als een bedreiging voor hun werk zien. Deze medewerkers hebben veel invloed gedurende de ontwikkeling. De organisatie is al jaren gewend aan de analyses zoals deze handmatig door de medewerkers uitgevoerd is, en nemen dat voor waar aan. Het is hierbij van belang dat er goede afstemming plaatsvindt met de analyse medewerkers om het resultaat van de analyse op waarheid te toetsen.

Tegelijkertijd is het mogelijk dat deze medewerkers niet of nauwelijks bereid zijn om mee te werken aangezien zij de komst van een analyse tooling als bedreiging zien voor hun eigen werkzaamheden. Het is van belang om deze personen tevreden te houden gedurende het traject. Tevens is het handig om deze personen input en feedback te laten geven op gerealiseerd werk. Hierbij wel gelet op de mogelijk negatieve benadering.

7.3 Stakeholder Matrix algemeen

Tot slot is er ook een stakeholder matrix opgesteld waarin wordt aangegeven welke externe belangen er mogelijk ook een rol spelen. Onderstaande drie groepen met belangen zijn in de stakeholder matrix terug te vinden.

- Algemeen belang gekenmerkt met A-
- Centric gekenmerkt met C-
- Strukton gekenmerkt met S-



Figuur 11, Algemene stakeholdermatrix

Onderstaand toelichting op de algemene belangen en de belangen van Centric. De belangen van Strukton zijn terug te vinden in hoofdstuk 7.2 Stakeholder Matrix Strukton Rolling Stock. (MindTools, 2016)

A-Media

Machine learning, big data, business intelligence, predictive maintenance en gerelateerde zaken zijn termen die in de hedendaagse vakliteratuur veel aandacht trekken. Zowel geslaagde als falende projecten komen hierbij aan bod.

Om die reden is het ook van belang dat de resultaten die in de buitenwereld gecommuniceerd worden positief van aard zijn. Wanneer dit niet zo is, kan de media door een negatief artikel voor veel negatieve invloeden van buitenaf zorgen. Door vooral voor de buitenwereld te benadrukken wat wel goed gaat, en waarin vooruitstrevend gehandeld is, kan de positieve publiciteit voor nieuwe belangstelling zorgen.

A-Toekomstige Klant

Toekomstige klanten van Centric, of een andere IT-organisatie hebben veel interesse in de ontwikkeling en de mogelijkheden van de Cortana Intelligence Suite, maar nauwelijks invloed. Voor toekomstige klanten is het ontwikkelen van het kennisniveau van de IT-dienstverlener van belang. Wanneer deze organisatie klant wordt en om dezelfde of een vergelijkbare dienstverlening vraagt, is de ervaring, kennis en expertise aanwezig bij de IT-dienstverlener.

Ook de toekomstige klanten van de concurrent hebben belang bij de ontwikkeling en mogelijkheden van de Cortana Intelligence Suite. Wanneer een andere organisatie een fout of probleem tegen komt, meldt deze organisatie het probleem bij de leverancier van de technology. In dit geval bij Microsoft. Op basis hiervan wordt het probleem opgelost over verder doorontwikkeld. Dit is voor de toekomstige klant van de concurrent van belang, zodat de mogelijkheden verreikt worden, en de al geconstateerde problemen opgelost zijn.

A-Microsoft

Microsoft heeft baad bij early adopters en de ervaringen die zij opdoen. Op basis van de informatie en ervaringen van de early adopters, kan een dienstverlening van Microsoft verder doorontwikkeld worden. Hierdoor heeft Microsoft wel interesse in de ervaringen en toepassingen, maar geen invloed.

De interesse van Microsoft is daarbij wel gering. De ervaringen van een individu zijn vele malen minder interessant dan die van de een grote groep gebruikers. Op basis hiervan is Microsoft in de neutrale categorie ingedeeld.

C-Stagiair

Voor de afstudeer stagiair is de toepassing van de Cortana Intelligence Suite van groot belang voor het slagen of falen van de afstudeerstage. Hierin is hij voorstander van toepassing van de Cortana Intelligence Suite, ook voor externe partijen. Het belang wat hij hierbij heeft, is het leren van ervaringen van anderen. Wanneer externe partijen vragen stellen of publicaties naar buiten brengen met betrekking tot de Cortana Intelligence Suite, kan dit input gevend zijn voor de afstudeerstage.

C-Begeleiders

De begeleiders van de afstudeer stagiair behartigen verschillende belangen. Het organisatiebelang en de positie van de organisatie op de markt speelt een rol, maar

tegelijkertijd ook de ondersteuning van de afstudeer stagiair bij de invulling van de afstudeeropdracht. Hierin zijn de begeleiders supporters en steun gevend.

C-Sales en C-Bedenker opdracht

De ervaring die opgedaan wordt met de Cortana Intelligence Suite is van belang voor de sales afdeling van Centric Advanced Solutions. Wanneer de ontwikkelde oplossing met de Cortana Intelligence Suite voldoet in de behoeften van Strukton R.S., en aansluiting vindt in de industriële sector, ligt hier een kans voor Centric Advanced Solutions om meerdere projecten binnen te halen waarbij de Cortana Intelligence Suite of delen hiervan ingezet wordt. Hiervoor dienen de accountmanagers van de sales afdeling de onderzochte mogelijkheden mee te nemen in hun portfolio.

Dit geldt eveneens voor de bedenker van deze afstudeeropdracht. Dhr. Will Loermans is accountmanager van Centric Solution Engineering waar Advanced Solutions mee verbonden is.

De sales afdeling van Centric Advanced Solutions zijn supporter van het idee. Een van de accountmanagers is tevens bedenker van de afstudeeropdracht. Salesmedewerkers kunnen geregeld een bonus krijgen bij goede resultaten. Hiermee kan een extra product wat aansluit bij de markt voor de salesmedewerkers een drive zijn voor actieve benadering van de spelers in deze markt.

C-AS (Advanced Solutions)

Voor Centric Advanced Solutions breed is dit project een ontwikkeling van kennis. De kennis wordt ontwikkeld op de inhoud, de mogelijkheden en toepassingen van de Cortana Intelligence Suite. De opgedane kennis wordt gedeeld, en draagt bij aan het kennis en expertise niveau van de medewerkers bij Advanced Solutions.

Deze kennis en ervaringen kunnen ingezet worden bij de latere toepassingen van de Cortana Intelligence Suite voor klanten in zowel de industriële sector en mogelijk ook andere sectoren.

C-Competence Centers

Binnen Centric zijn er verschillende Competence centers waar kennis op verschillende gebieden organisatie breed gedeeld wordt. Zo is er een Competence Center big data, business intelligence en Microsoft Azure die alle drie raakvlak hebben met het onderzoek naar de toepassing van de Cortana Intelligence Suite in de industriële sector.

Na afronding van de opdracht wordt de opgedane kennis met alle drie de Competence Centers binnen Centric gedeeld. Deze kennis kan bijdragen aan het algehele kennisniveau van eenieder die actief deelneemt aan deze Competence Centers.

7.4 Conclusie

Voor Strukton zijn er meerdere belangen die behartigd worden met analyse van de remote monitoring data. Grofweg is dit op te splitsen in de volgende categorieën:

- Productverbetering op basis van geregistreerde fouten uit het verleden.
- Voorraadbeheer van onderdelen op basis van informatie over kwetsbaarheid en levensduur van de onderdelen.
- Doorontwikkeling van de software en treinbesturingssystemen op basis van fout analyse.

- Gebruik hedendaagse technologie wat bijdraagt aan de positionering van Strukton op de markt.

Buiten de belangen voor Strukton Rolling Stock, spelen er ook nog belangen mee voor zowel Centric als algemeen belang. Onderstaand de overige belangen.

Algemeen belang:

- Media, toepassing van nieuwe technologieën draagt bij aan publicaties over de mogelijkheden van de technologie mits deze informatie naar buiten wordt gebracht.
- Experimenteren met de technologie draagt bij aan het kennisniveau. Dit is van belang voor mogelijk toekomstige klanten waarbij vooraf de mogelijkheden en onmogelijkheden duidelijk zijn en toepassingen sneller mogelijk zijn.
- Experimenteren met de technologie is eveneens van belang voor Microsoft, de ontwikkelaar van de technologie om van de fouten en problemen die optreden te leren en door te ontwikkelen.

Centric belang:

- Opbouw van kennis en behalen van diploma met afronding van de afstudeeropdracht voor de stagiair.
- Begeleiders van de afstudeer stagiair voor de opbouw van kennis en het behartigen van de belangen van de stagiair.
- Sales medewerkers binnen Centric voor het op de markt brengen van een nieuw vooruitstrevend product waar uiteenlopende toepassingsmogelijkheden mee zijn.
- Afdeling Advanced Solutions voor het verkennen van de mogelijkheden, de kennisdeling en ervaring opdoen met de Cortana Intelligence Suite.
- Verschillende Competence Centers binnen Centric voor verkenning van de mogelijkheden en deling van kennis.

8 Veld: Betekenis data

De registraties van Strukton Rolling Stock die worden aangemaakt wanneer er een fout optreedt in een locomotief, hebben vele variabelen in zich en een vastgesteld format. Om deze gegevens te vertalen naar informatie, wordt in dit hoofdstuk toegelicht hoe de data opgebouwd is, en welke variabelen hierin zitten. Ook wel data-analyse genoemd.

In hoofdstuk 8.1 wordt de methode en werkwijze besproken die is toegepast tijdens de data-analyse. Hierop volgend zijn in hoofdstuk 8.2 de resultaten van de data-analyse toegelicht. Tot slot in hoofdstuk 8.3 de Conclusie van de data-analyse.

8.1 Methode en werkwijze

Gedurende de volledige periode van het afstuderen, ben ik wekelijks twee dagen bij Strukton Rolling Stock in Alblasterdam aanwezig geweest. Tijdens deze periode hebben verschillende personen binnen Strukton mij bijgepraat over de data die verzameld wordt met de remote monitoring box. De remote monitoring box is de hardwarecomponent die ervoor zorgt dat de registratie data wordt verzameld, opgeslagen en wanneer mogelijk doorgestuurd naar de backoffice van Strukton

In de data-analyse worden de variabelen behandeld die in de registraties aanwezig zijn. Onderstaand wordt de methode en werkwijze met betrekking tot de vaststelling van de betekenis van de gegevens uitgewerkt. Aansluitend hierop is het sterschema, ook wel datamodel, ingericht om de registraties in een relationele database weg te schrijven. De methode en werkwijze die hierbij zijn toegepast zijn in dit hoofdstuk eveneens uitgewerkt.

Vaststellen betekenis gegevens

Om de inhoud van de registraties te analyseren, is er met dhr. Marcel van der Linden, software developer bij Strukton R.S., stap voor stap een registratie ontleed. Hierbij hebben we gezamenlijk ieder onderdeel uit de data bekeken om de betekenis hiervan te verduidelijken.

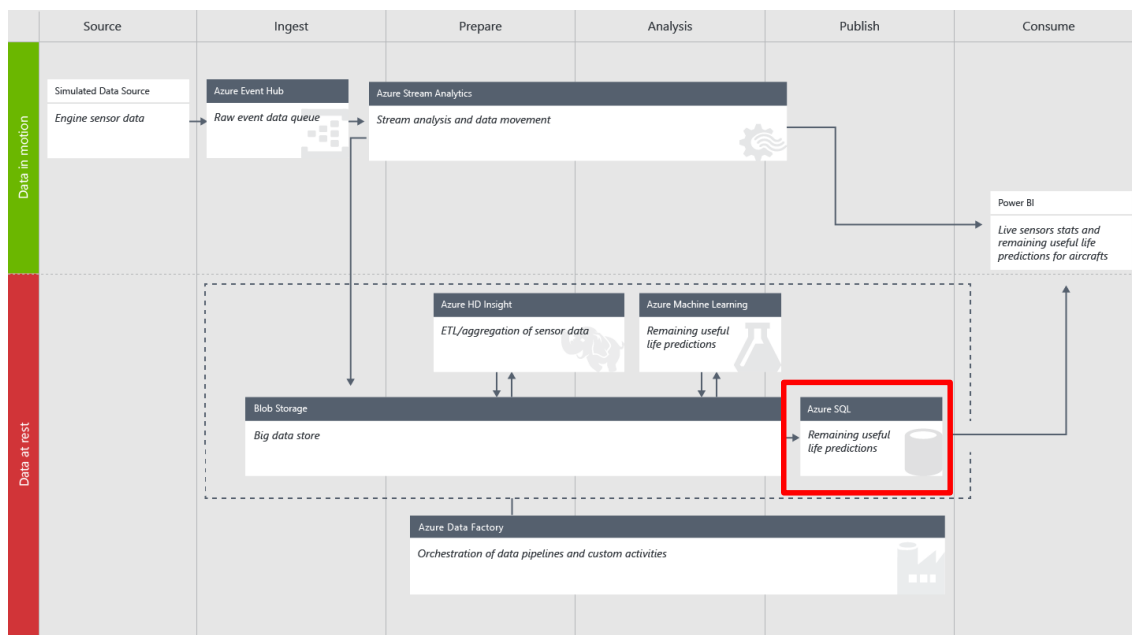
Aansluitend hierop is er gezamenlijk met Marcel van der Linden een gesprek geweest met dhr. Marius Klein. Marius Klein heeft vanuit zijn functie de huidige analyse tooling SCARP ontwikkeld. In deze tool worden dezelfde XML-registratiebestanden gebruikt. Hierdoor weet Marius Klein exact de opbouw van de XML-registratie bestanden, hoe data is opgebouwd en hoe eventuele waarden berekend kunnen worden.

De werkwijze die gehanteerd is tijdens het gesprek met Marcel van der Linden, is niet toe te kennen aan een specifieke methode of techniek. Gezamenlijk is er tijdens dit gesprek een registratie bestand van het begin tot het eind doorlopen op gestructureerde wijze.

Het gesprek met Marius Klein is ontstaan na een aantal vraagstukken die zijn gevolgd uit de sessie met Marcel van der Linden. Dit kan beoordeeld worden als gestructureerd kwalitatief diepte-interview omdat er specifieke vragen gesteld zijn en die kenmerken zijn van een gestructureerd kwalitatief diepte-interview. Het voordeel van een diepte-interview is dat er gemakkelijk doorgevraagd kan worden en op specifieke gevallen ingegaan kan worden. (Alles over marktonderzoek, n.d.)

Opstellen datamodel

Om alle registraties weg te kunnen schrijven in een relationele database, is een ster-schema opgesteld. Kijkend naar de schematische weergave van de Cortana Intelligence Suite, moet de data uiteindelijk weggeschreven worden in een SQL-database voordat deze data gevisualiseerd kan worden met Microsoft Power BI. Dit is te zien in Figuur 12, Schematische weergave Cortana Intelligence Suite. In deze procesflow is de Azure SQL met rood gemarkeerd in de stap voordat de data naar de visualisatie tooling Power BI gaat.



Figuur 12, Schematische weergave Cortana Intelligence Suite

Het datamodel geeft structuur aan de data en maakt het mogelijk om bevestigingen op de data effectiever en sneller uit te voeren. Zo is het niet effectief om alle data in dezelfde tabel weg te schrijven. Hierbij ontstaan veel dubbele gegevens en overhead in de data. Op den duur vervuilen de dubbele waarden de database en gaat dit ten koste van de performance.

Tevens is het gemakkelijker om een filter toe te passen in een bevestiging van de database wanneer er gebruik wordt gemaakt van een datamodel. Ook hier komt dit de performance van de bevestiging ten goede.

Om het datamodel op te stellen is gebruik gemaakt van het feit en dimensie model. Deze keuze is gemaakt omdat deze methode veel gebruikt wordt voor business intelligence toepassingen, waar de Cortana Intelligence Suite veel raakvlak mee heeft. In dit model wordt onderscheid gemaakt tussen feiten en dimensies.

Een feittabel bestaat vaak uit meetwaarden en heeft de variabele waarden in zich. Een dimensietabel is een tabel waarin de objecten en vaste waarden uit de data weggeschreven zijn. Vanuit de feittabellen wordt vaak gerefereerd naar de dimensietabellen om enkele van de vaste waarden op te vragen.

(About Tech - Mike Chapple, database expert, 2014)

8.2 Data-analyse

In dit hoofdstuk worden de resultaten van de data-analyse beschreven. Om de betekenis van de data goed in kaart te brengen, is in onderstaande tabel elke variabele uit de data te vinden met de daarbij behorende omschrijving. Aansluitend hierop worden de variabelen uitgewerkt die berekend dienen te worden op basis van de aanwezige variabelen. Tot slot is er een datamodel waarin de data effectief ingedeeld wordt voor de inrichting een relationele database.

De data van Strukton Rolling Stock ziet erin .XML-format als volgt uit:

```
<?xml version="1.0"?>
<SavedRegistrationList xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xmlns:xsd="http://www.w3.org/2001/XMLSchema">
  <ProjectID>0000388018</ProjectID>
  <BoardID>Dcu3</BoardID>
  <BoardSerialNumber>
    <Value xmlns="http://StruktonSystems/SDU">000034</Value>
  </BoardSerialNumber>
  <Records>
    <RecordNumber xmlns="http://StruktonSystems/SDU">1</RecordNumber>
    <NumberOfBlocks xmlns="http://StruktonSystems/SDU">450</NumberOfBlocks>
    <Length xmlns="http://StruktonSystems/SDU">89808</Length>
    <Version xmlns="http://StruktonSystems/SDU">
      <Major>235</Major>
      <Minor>7</Minor>
      <Temp>0</Temp>
    </Version>
    <DateTime xmlns="http://StruktonSystems/SDU">2013-04-11T13:33:09</DateTime>
    <Faults xmlns="http://StruktonSystems/SDU">
      <Text>ID: 68 --&gt; SLG1_06_BVCBofWidDis</Text>
    </Faults>
    <TriggerFaults xmlns="http://StruktonSystems/SDU">
      <Text>ID: 68 --&gt; SLG1_06_BVCBofWidDis</Text>
    </TriggerFaults>
    <Mask xmlns="http://StruktonSystems/SDU">268435455</Mask>
    <FastFrequency xmlns="http://StruktonSystems/SDU">10000</FastFrequency>
    <FastSize xmlns="http://StruktonSystems/SDU">54000</FastSize>
    <SlowFrequency xmlns="http://StruktonSystems/SDU">10</SlowFrequency>
    <SlowSize xmlns="http://StruktonSystems/SDU">35640</SlowSize>
    <TimeAfterHistory xmlns="http://StruktonSystems/SDU">0.0500000045</TimeAfterHistory>
    <Signals xmlns="http://StruktonSystems/SDU">
      <AnalogSignal>
        <Name>F916_IFilt_difxtime</Name>
        <Description>F916_IFilt_difxtime --&gt; t.b.d.</Description>
        <Data>
          <Sample xsi:type="FloatValue">
            <Value>0</Value>
          </Sample>
          <Sample xsi:type="FloatValue">
            <Value>0</Value>
          </Sample>
          <Sample xsi:type="FloatValue">
            <Value>0</Value>
          </Sample>
        </Data>
      </AnalogSignal>
    </Signals>
  </Records>
</SavedRegistrationList>
```

Figuur 13, Voorbeeld .XML-data Strukton Rolling Stock

Variabelen in de data

Onderstaande tabel bevat alle gegevens die in de registraties te vinden zijn. Bij ieder gegeven wordt een korte omschrijving gegeven. De omschrijvingen zijn voortgekomen uit het gesprek met Marcel van der Linden en Marius Klein van Strukton Rolling Stock.

Gegeven in registratie: **Omschrijving:**

ProjectID	Met het Project ID wordt een unieke waarde aangegeven van de organisatie die de meting uitvoert. In dit geval is deze waarde bij alle registraties hetzelfde.
-----------	---

BoardID	Het Board ID is een waarde die aangeeft op welke hardwarecomponent de meting is uitgevoerd.
BoardSerialNumber	Het serienummer van het board wat is aangegeven bij BoardID.
RecordsNumber	Aantal registraties aanwezig in het bestand.
NumberOfBlocks	Aantal mogelijke meetwaarden in registratie.
Length	Totale lengte van de registratie in tijd. Bevat de som in tijd van de triggerfault, slowsize en fastsize.
Version	Versie van de software die op de hardware component is geïnstalleerd.
Major	Versienummer voor de punt van de SCARP config die aan de data wordt toegevoegd op de backoffice om de data betekenis te geven.
Minor	Versienummer na de punt van de SCARP config die aan de data wordt toegevoegd op de backoffice om de data betekenis te geven.
Temp	Derde (niet tot nauwelijks gebruikte) variabele van het versienummer van de SCARP config.
DateTime	Datum en tijd waarop de meting is afgerond en opgeslagen op de remote monitoring box.
Faults	Namen van de fouten die zijn opgetreden tijdens de registratie.
Triggerfaults	De namen van de fout(en) die als trigger hebben gefungeerd en tot de meting hebben geleid. De datum en tijd van de trigger is te berekenen door $DateTime - TimeAfterHistory$.
Mask	De MAS is ene bit masker dat aangeeft welke analoge signalen er gevuld zijn. Wanneer een signaal niet wordt gebruikt of door het board wordt afgekeurd, dan staat er een 0 bij het bijbehorende bit.
FastFrequency	Samplerate, ook wel frequentie van de metingen na het moment van de triggerfault. Deze periode heeft een hogere samplerate dan de periode voor de triggerfault. De tijd per sample is te berekenen door $1/FastFrequency$.
FastSize	Omvang van de FastFrequency.
SlowFrequency	Samplerate, ook wel frequentie van metingen voor het moment van de triggerfault. Deze periode heeft een lagere samplerate dan de periode na de triggerfault. De tijd per sample is te berekenen door $1/SlowFrequency$.
SlowSize	Omvang van de SlowFrequency.
TimeAfterHistory	De duur van de periode tussen het moment dat de trigger plaats vindt, en de afsluiting van de registratie.
AnalogSignal	
Name	Naam van het gemeten signaal.
Description	Omschrijving van het gemeten signaal.
Sample	Meetwaarde van het analoge signaal. Dit is een variabele waarde.
DigitalSignal	

Name	Naam van het gemeten signaal.
Description	Omschrijving van het gemeten signaal.
Sample	Meetwaarde van het digitale signaal. Bij een digitaal signaal is dit altijd 0 of 1.
InstallationNumber	Versie van de software die draait op de hardwarecomponent waar de meting op uitgevoerd is.
iButtonData	Wordt afhankelijk van het project anders ingezet.

Tabel 2, Betekenis variabelen Strukton remote monitoring data**Berekende variabelen**

Op basis van de gegevens die aanwezig zijn, zijn er ook enkele waarden die berekend moeten worden. Tijdens de sessie met Marcel van der Linden zijn deze berekende waarden aan de orde gekomen. Marius Klein heeft ons vervolgens voorzien van de juiste informatie om deze variabelen te kunnen berekenen.

Onderstaand overzicht toont enkele van de waarden die berekend moeten worden met daarbij de juiste variabelen en formules. Hierbij wordt ook een korte omschrijving gegeven waarom deze waarde van toegevoegde waarde kan zijn om aan de dataset toe te voegen.

Datum en tijd van de Triggerfault

De datum en tijd waarop de triggerfault heeft plaatsgevonden kan van toegevoegde waarde zijn om te kunnen achterhalen wat de oorzaak van de fout is die heeft geleid tot de registratie.

$$\text{DateTime Triggerfault} = \text{DateTime} - \text{TimeAfterHistory}$$

Doorlooptijd van de SlowSize en de FastSize

De doorlooptijd van de SlowSize en FastSize, ook wel hoe lang wordt de meting uitgevoerd wat heeft geleid tot de meetwaarde in de sample, kan van belang zijn voor de analyse wat tot de betreffende fout in de registratie heeft geleid. Deze waarde wordt in seconden aangegeven.

$$\text{Doorlooptijd Samples SlowSize} = \frac{\text{SlowSize}}{36} \times \frac{1}{\text{SlowFrequency}}$$

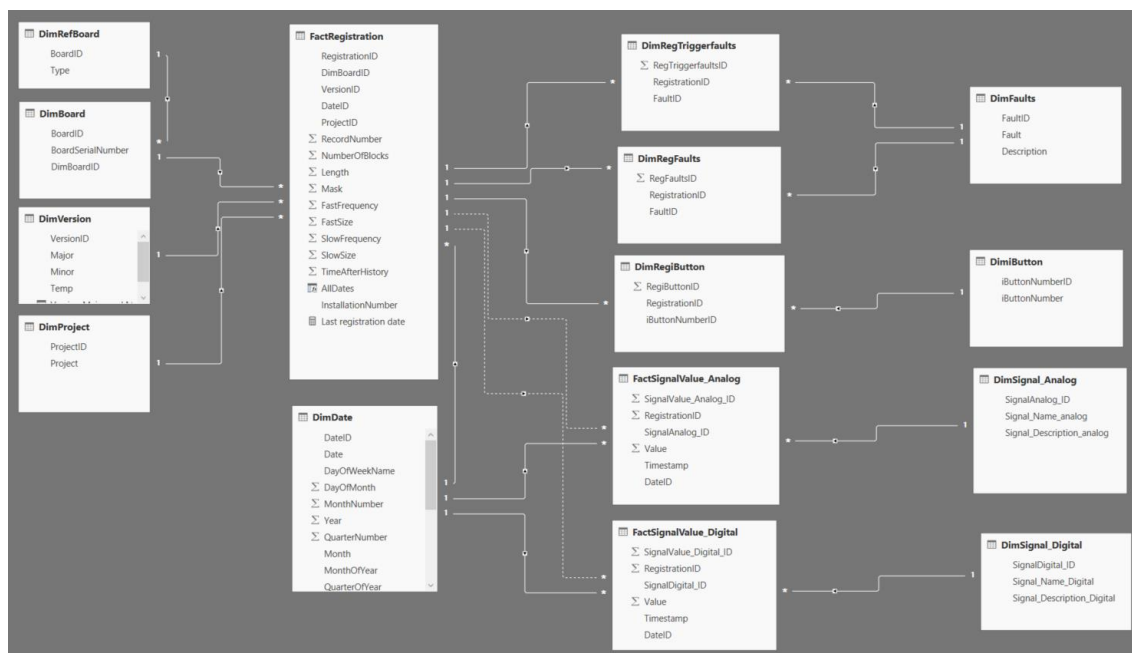
$$\text{Doorlooptijd Samples FastSize} = \frac{\text{FastSize}}{36} \times \frac{1}{\text{FastFrequency}}$$

In bovenstaande berekeningen zijn 1 en 36 vaste waarden.

Om te lokaliseren hoeveel samples er binnen de SlowSize en hoeveel er binnen de FastSize vallen, kan er ook gekeken worden naar de verandering in stap grootte op basis van de timestamp van de meetwaarde.

Datamodel

Zoals in hoofdstuk 8.1 bij methode en werkwijze is toegelicht, is het database technisch niet verstandig om één tabel aan te maken waarin alle data wordt weggeschreven. Hiervoor is onderstaand feit/dimensie model opgesteld. Onderstaand datamodel is de definitieve versie zoals dit gerealiseerd is voor de operationele Power BI dashboard. De doorontwikkelingen van het datamodel zijn naar aanleiding van aanvullende data of fouten in de relatie tussen de feiten en dimensies.



Figuur 14, Datamodel Azure SQL-database

8.3 Conclusie

Op basis van de data valt te concluderen dat domeinkennis van groot belang is in de situatie van Strukton Rolling Stock. Zonder kennis van de gegevens en de betekenis hiervan, zijn er geen relevante visualisaties en nieuwe inzichten te genereren.

Deze conclusie kan getrokken worden omdat er naar eigen interpretatie slechts enkele van de gegevens uitgelegd kunnen worden. Overige gegevens zijn met behulp van zowel Marcel van der Linden als Marius Klein ingevuld en uitgewerkt.

De resultaten van de data betekenis zijn benodigd als input om de Veld: Informatiebehoefte Strukton Rolling Stock zoals beschreven in hoofdstuk 9 te vertalen naar functionele oplossingen. Deze functionele oplossingen moeten uiteindelijk aansluiten bij de beleving en behoeften van de (eind) gebruikers en gevisualiseerd worden in Power BI zoals beschreven wordt in hoofdstuk 9.

9 Veld: Informatiebehoefte Strukton Rolling Stock

Voorafgaand aan de ontwikkeling van visualisaties op basis van de data van Strukton Rolling Stock, is geïnterviewd waar de informatiebehoefte ligt. De wijze waarop deze informatiebehoefte is vastgesteld is beschreven in hoofdstuk 9.1. Hierop volgend zijn de resultaten van deze werkwijze toegelicht in hoofdstuk 9.2. Tot slot wordt de conclusie met betrekking tot de informatiebehoeften gegeven in hoofdstuk 9.3.

9.1 Methode en werkwijze

Om de behoeften van de werknemers bij Strukton Rolling Stock inzichtelijk te maken, is er gebruik gemaakt van interviews met de belangrijkste stakeholders. In eerste instantie zijn hierbij de directbetrokkenen bij de afstudeeropdracht geïnterviewd. Deze interviews hadden een kwalitatief karakter en ongestructureerd van aard. Hierdoor was het gemakkelijker om door te vragen op basis van de gegeven antwoorden. Om geen belangrijke onderwerpen tijdens de interviews te vergeten, is gebruik gemaakt van een korte lijst met onderwerpen die aan bod moeten komen gedurende het interview. Ook is iedere geïnterviewde gevraagd naar eventueel andere relevante personen om te interviewen.

(Alex Feijt - Research Company, 2013)

De input uit deze interviews heeft het startpunt gevormd voor de informatiebehoeften waarvoor in de PoC naar een antwoord is gezocht. De resultaten die hieruit voortgekomen zijn, zijn gepresenteerd aan een grotere groep potentiële gebruikers in de vorm van een demo. Tijdens deze demo is er aan de groep gebruikers gevraagd om aanvullende informatiebehoeften feedback op basis van het gevisualiseerde. De toelichting op de demo wordt nader beschreven in hoofdstuk 11.

Om de vastgestelde informatiebehoeften te prioriteren, is er gebruik gemaakt van de MoSCoW-methode. Bij de MoSCoW-methode wordt er onderscheid gemaakt in vier niveaus. De volgende niveaus worden hierin onderscheiden:

- **Must-have**, vereist onderdeel om te kunnen spreken van een werkbaar product.
- **Should-have**, hoge prioriteit maar niet vereist voor een werkbaar product.
- **Could-have**, optioneel, wordt alleen meegenomen als er tijd over is.
- **Would-have**, geen prioriteit, kan eventueel in de toekomst worden overwogen.

(Marketingtermen, n.d.)

De MoSCoW-methode is toegepast om de informatiebehoeften en aanvullingen naar aanleiding van de demo te prioriteren.

9.2 Vastgestelde informatiebehoefte

In eerste instantie zijn er een aantal standaard vragen opgesteld als startpunt. Deze vragen zijn opgesteld in samenwerking met Marcel van der Linden en Han Geerligs van Strukton Rolling Stock. De gestelde vragen zijn als startpunt gekozen om aansluitend hierop nieuwe vragen te formuleren. Deze vragen zijn opgesteld op basis van de gegevens die in de data aanwezig zijn. Deze gegevens en de betekenis hiervan is terug te lezen in hoofdstuk 8. Onderstaande opsomming bevat de vragen zoals deze vooraf zijn opgesteld.

- Welke fout treedt er het meeste op?
- Welke trigger leidt tot de meeste registraties, en hoe vaak?

- Is er een tijdsrange of periode waarin de meeste fouten optreden?
- Hoeveel fouten zijn er per board?
- Hoeveel fouten zijn er per locomotief (installatienummer)?
- Welke versie of serienummer van het board vertoont het meeste fouten?

Naast bovenstaande specifieke vragen, is het de bedoeling dat de data van de registratie in het dashboard, gefilterd kan worden op de volgende elementen:

- Locomotief;
- Board;
- Versienummer;
- Tijd.

De bovenstaande informatie heeft geleid tot een eerste versie van het dashboard in Power BI. Dit is in eerste instantie met een demo dataset gedaan. De dataset die voor de demo gebruikt is, was behoorlijk beperkt en handmatig samengesteld. Hierdoor was het niet mogelijk om op ieder bovengenoemd gebied visualisaties te tonen. Het Power BI-dashboard met de demo dataset wordt nader toegelicht in hoofdstuk 11.

Doorontwikkeling

Naar aanleiding van de demo zijn verschillende aanvullingen en verbeteringen aangedragen door stakeholders (zie hoofdstuk 7.2) bij Strukton Rolling Stock. Deze aanvullingen bestonden uit het combineren van andere velden om verbanden te ontdekken, maar ook het combineren van andere databronnen. Onderstaande databronnen zijn als aanvulling genoemd:

- Verkoopgegevens, om de verhouding tussen de vloot en de fouten inzichtelijk te maken.
- Onderhoudsgegevens, om te kijken wanneer welke locomotief onderhoud heeft gehad en de relatie te leggen met de fouten.
- Locatiegegevens, van de registraties om mogelijke oorzaken te achterhalen van de fouten.

Aansluitend hierop zijn er nog de mogelijkheden waarbij open-data gecombineerd wordt. Deze mogelijkheden zijn niet aan de orde gekomen tijdens de demo, maar deze ideeën zijn op later moment ontstaan. De volgende open-databronnen kunnen gecombineerd worden met de data van Strukton.

- Meteorologische data, om verbanden en de afhankelijkheid van weersomstandigheden met de fouten inzichtelijk te maken. Om dit op een relevante wijze toe te passen, zijn locatiegegevens vereist.

Wanneer de vastgestelde informatiebehoeften ingedeeld worden in de MoSCoW-methode, ontstaat de volgende verdeling:

	Must-have:	Should-have:	Could-have:	Would-have:
1	Welke fout treed er het meeste op?	Is er een tijdsrange of periode waarin de meeste fouten optreden.	Locatiegegevens van de registraties om mogelijke oorzaken te achterhalen van de fouten.	Verkoopgegevens om de verhouding tussen de vloot en de fouten inzichtelijk te maken.

2	Welke trigger leidt tot de meeste registraties, en hoeveel zijn dat er?	Patroonherkenning en fouten aan zien komen op basis van de signalen bijbehorend bij de registraties.	Onderhoudsgegevens om te kijken wanneer welke locomotief onderhoud heeft gehad en de relatie te leggen met de fouten.
3	Hoeveel fouten zijn er per board?		Meteorologische data om verbanden en de afhankelijkheid van weersomstandigheden met de fouten inzichtelijk te maken.
4	Hoeveel fouten zijn er per locomotief (installatienummer)?		
5	Welke versie of serienummer van het board vertoont het meeste fouten?		

Tabel 3, MoSCoW indeling informatiebehoeften

Gedurende de uitvoering van de proof of concept is er enige vertraging ontstaan op technisch vlak. Hierdoor is besloten patroonherkenning uit de kolom Should-have verschoven naar Could-have. In verband met de beperkte tijdspanne is het voorspellen op basis van de beschikbare data buiten de scope van de PoC geplaatst.

9.3 Conclusie

De inventarisatie van de informatiebehoefte is voor de proof-of-concept met behulp van interviews achterhaald. Deze informatiebehoefte is vervolgens omgezet in een demo dashboard welke wordt getoond in hoofdstuk 11. Op basis van deze demo is er om feedback gevraagd bij de uiteindelijke gebruikersgroep, waarmee de aanvullende informatiebehoefte kenbaar is gemaakt voor de operationele versie van het dashboard.

Op basis van de operationele versie is er wederom om feedback gevraagd bij de gebruikersgroep waaruit enkele kleine aanpassingen zijn voort gekomen. Na deze kleine bewerkingen is de proof-of-concept fase afgerond en is een basis gevormd voor doorontwikkeling op langere termijn.

10 Werkplaats: Elementen Cortana Intelligence Suite

De Cortana Intelligence Suite is opgebouwd uit meerdere elementen uit de Microsoft Azure Stack. In dit hoofdstuk wordt beschreven welke elementen uit de suite gebruikt zijn gedurende de proof-of-concept om tot het uiteindelijke resultaat welke worden beschreven in hoofdstuk 11 te komen. Hiervoor wordt allereerst in hoofdstuk 10.1 beschreven welke methode en werkwijze is toegepast. Vervolgens wordt in hoofdstuk 10.2 toegelicht welke elementen in de Cortana Intelligence solution template zitten, en wat deze solution template exact is. Aansluitend hierop wordt in hoofdstuk 10.3 beschreven welke elementen uit de suite zijn gebruikt voor de proof-of-concept.

10.1 Methode en werkwijze

Om de toepassing van de Cortana Intelligence Suite te toetsen, is er gebruik gemaakt van experimenteel onderzoek. Binnen de kaders van experimenteel onderzoek is onderscheid te maken in verschillende typen. Het type wat het beste past bij de onderzoek casus is een quasi-experimenteel onderzoek. Dit type onderzoek wordt veelal gebruikt bij praktijkgericht onderzoek. Aangezien er voor het onderzoek naar toepassing van de Cortana Intelligence Suite gebruik wordt gemaakt van een PoC aan de hand van de situatie bij Strukton, spreken we van een praktijkcasus waarbij de toepassing van de suite getest wordt. (Wetenschap.infonu.nl - Rebecca Hoogers, 2011)

Bij experimenteel onderzoek wordt gebruik gemaakt van een hypothese die getoetst wordt. De uitkomst van het experiment bewijst of ontkracht vervolgens de hypothese. De onderzoeksvraag zoals deze is opgesteld en terug te lezen is in hoofdstuk 3.6, kan worden geformuleerd als hypothese.

De hypothese die getoetst wordt tijdens het experimenteel onderzoek is als volgt:

“De Cortana Intelligence Suite vormt een set aan cloud diensten die samen een totaaloplossing bieden voor de industriële sector.”

Om deze hypothese te toetsen, wordt de proof-of-concept voor Strukton Rolling Stock gebruikt. De data van Strukton wordt als uitgangspunt genomen om te werking van de Cortana Intelligence Suite te toetsen.

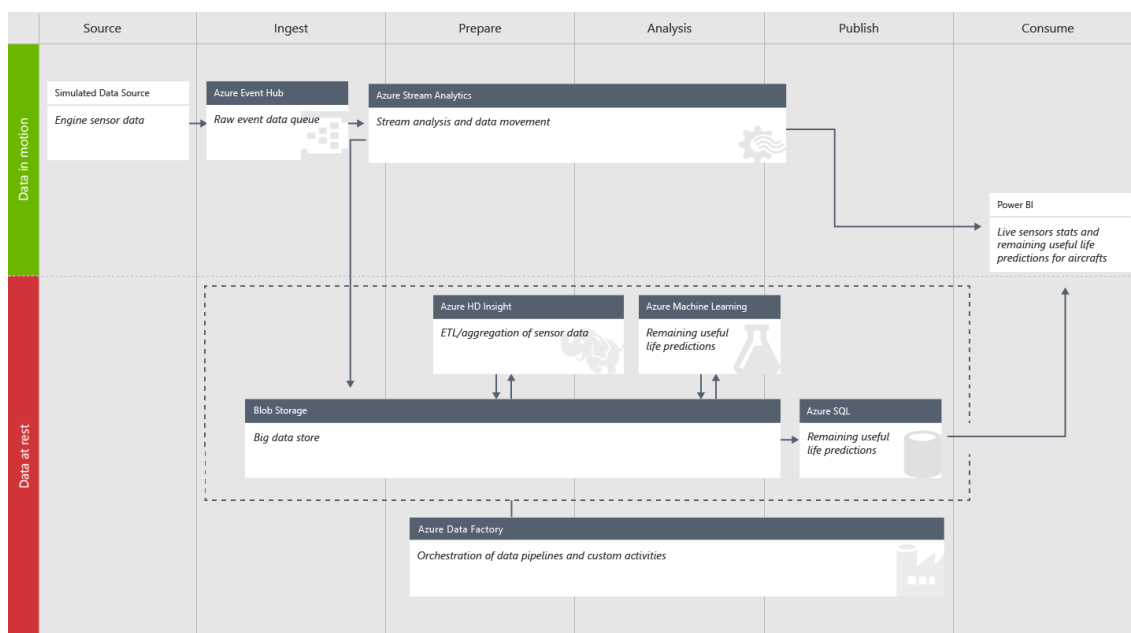
10.2 Solution Template

Microsoft biedt naast de diensten uit de Cortana Intelligence Suite ook een gallery waarin elementen uit de suite als experimenten en template aangeboden worden. Microsoft biedt solutions templates, experimenten, machine learning api's, documentatie en tutorials aan via de website van de Cortana Intelligence Gallery.

In deze gallery worden een aantal templates van complete ingerichte toepassingen van de Cortana Intelligence Suite aangeboden. Deze templates zijn te gebruiken als voorbeeld, maar ook als startpunt voor de proof-of-concept. Bij iedere solution template wordt documentatie om de template aan te kunnen passen naar de eigen behoeften en situatie geleverd.

De gebruikte template uit de Cortana Intelligence Gallery betreft een template waar op basis van diverse meetgegevens uit een vliegtuigmotor berekend wordt wat de levensduur van de motor nog betreft. Hierbij worden de volgende onderdelen uit de Cortana Intelligence Suite gebruikt:

- Azure Event Hub;
- Azure Stream Analytics;
- Azure HD Insight;
- Azure Machine Learning;
- Azure Blob Storage;
- Azure SQL;
- Azure Data Factory;
- Power BI.



Figuur 15, Cortana Intelligence Suite - Solution Template

Bovenstaande onderdelen vormen een aaneenschakeling van Azure elementen die samen de totaaloplossing en de status van de vliegtuigmotor tonen in Power BI.

Na het deployen inclusief demo dataset uit de Cortana Intelligence Gallery, is er gekeken hoe de omgeving ingericht is. De bijgesloten documentatie van de solution template omschrijft hoe de template omgebouwd kan worden naar een omgeving waarbij de eigen data gebruikt kan worden. Voor de proof-of-concept met de data van Strukton is in eerste instantie gekozen om de template aan te passen naar de benodigde inrichting voor Strukton.

Aanpassing van de solution template

De Azure Event Hub, waarin de data binnenkomt en via de Azure Stream Analytics wegschrijft naar de Blob Storage, zorgde bij de aanpassing van de solution template voor problemen. Deze problemen hadden allereerst betrekking op de data import. Daarnaast ontstonden in een later stadium van de aanpassing van de solution template uitdagingen met betrekking tot de data conversie en het wegschrijven van de data naar een Azure SQL Database.

Data Import

De Azure Event Hub accepteert enkel .CSV en .JSON als inkomend bestandsformat. Aangezien de data van Strukton Rolling Stock als .XML weggeschreven wordt, kan deze data niet via de Azure Eventhub geïmporteerd worden in de procesflow zoals in Figuur 15. Dit probleem is aanleiding geweest een andere ingang voor de data te kiezen. Om de inkomende data in de toekomst via de Azure Event Hub te laten verlopen, kan op een later moment worden onderzocht.

Data wegschrijven naar Data Blob

De volgende optie was om de datadump van Strukton direct te uploaden naar de Blob Storage. Vanuit de blob storage kan Azure Machine Learning op de data worden toegepast en vervolgens via een gegevenspipeline in de Azure Data Factory weggeschreven worden naar de Azure SQL Database. Echter vereist de Azure SQL Database eveneens als de Even Hub een inkomend bestand met het format .CSV of .JSON.

Data conversie

Om het probleem rondom het bestandsformat op te lossen is er voor gekozen om met behulp van een programmeur een conversie te maken in de gegevenspipeline waarbij het .XML-format van Strukton omgezet wordt naar een .JSON. De keuze is gemaakt voor .JSON omdat dit bestand het minste overhead en opmaak bevat wat ervoor zorgt dat het bestand zo klein mogelijk is. Na verschillende tests blijkt de conversie naar .JSON erg veel tijd in beslag te nemen. Deze tijden liepen uiteen van enkele minuten tot meer dan 48 uur per bestand. Enkele aanpassingen en doorontwikkelingen op de conversiepipeline hebben het conversieproces versnelt, maar nog niet tot een acceptabele duur.

Data importeren in Azure SQL Database

De al geconverteerde bestanden zijn gebruikt om te testen of deze geïmporteerd kunnen worden in de Azure SQL Database. Echter bleek dat de geconverteerde bestanden om een nog onduidelijke reden niet geaccepteerd en geïmporteerd kunnen worden in het opgestelde datamodel die in Figuur 14 wordt weergegeven. Naar alle waarschijnlijkheid komt dit probleem voort uit het format bij de .JSON-bestanden die niet geheel volgens de vereiste structuur van de Azure omgeving is.

Aangezien er al de nodige tijd verstreken was om tot bovenstaande ontdekkingen te komen, is er besloten om naar een andere aanpak te kijken. De tijdsspanne die over was, was onvoldoende om verder te gaan op de ingeslagen weg, en in de beperkte tijd tot een werkende proof-of-concept te komen. Naar aanleiding van deze constatering is besloten om Azure Machine Learning voor de proof-of-concept buiten scope te plaatsen.

Om toch tot een werkend geheel te komen en voor Strukton inzichten te tonen op basis van de beschikbare data, is een andere opzet gekozen dan in de solution template gehanteerd wordt. De gebruikte elementen en de wijze waarom wordt beschreven in hoofdstuk 10.3.

10.3 Gebruikte elementen uit de Cortana Intelligence Suite

Om met de proof-of-concept voor Strukton Rolling Stock een werkende situatie op te leveren, is er gekozen voor een andere oplossing dan de solution template biedt. Hierbij zijn enkele elementen uit de solution template gebruikt om het uiteindelijke resultaat van de proof-of-concept te komen.

Power BI vereist een gestructureerde omgeving met data. In de solution template betreft dit de Azure SQL Database. Aangezien de .XML-bestanden van Strukton wel direct te importeren zijn in een SQL-database, is er voor gekozen om voorgaande stappen uit de solution template (zie Figuur 15) over te slaan en de data direct te importeren in de database.

Aanmaken Azure SQL-database

Er waren wel een aantal stappen nodig voordat de data in de juiste structuur in de Azure SQL Database weggeschreven konden worden. Zo moest vooraf het datamodel inclusief alle onderlinge relaties tussen de tabellen aangemaakt worden in de database. Om dit te realiseren, heb ik met het programma Microsoft Visual Studio verbinding gemaakt met de Azure SQL Server. Op de SQL Server is vervolgens een nieuwe database aangemaakt en daarin alle benodigde tabellen en bijbehorende relaties. Het aanmaken van de tabellen en relaties is gerealiseerd door gebruik te maken van SQL-statements.

Iedere benodigde tabel zoals dit is vastgesteld in het datamodel die te zien is in Figuur 14 is op gelijke wijze aangemaakt in de SQL Server database. De wijze waarop deze query's zijn opgebouwd is getoond in Figuur 16.

```

1 CREATE TABLE [dbo].[FactRegistration] (
2     [RegistrationID] INT IDENTITY (1, 1) NOT NULL,
3     [DimBoardID] INT NOT NULL,
4     [VersionID] INT NOT NULL,
5     [DateID] INT NOT NULL,
6     [ProjectID] INT NOT NULL,
7     [RecordNumber] INT NULL,
8     [NumberOfBlocks] INT NULL,
9     [Length] INT NOT NULL,
10    [Mask] INT NULL,
11    [FastFrequency] INT NOT NULL,
12    [FastSize] INT NOT NULL,
13    [SlowFrequency] INT NOT NULL,
14    [SlowSize] INT NOT NULL,
15    [TimeAfterHistory] FLOAT (53) NOT NULL,
16    [InstallationNumber] NCHAR (10) NULL,
17    PRIMARY KEY CLUSTERED ([RegistrationID] ASC),
18    CONSTRAINT [FK_FactRegistration_DimVersion] FOREIGN KEY ([VersionID]) REFERENCES [dbo].[DimVersion] ([VersionID]),
19    CONSTRAINT [FK_FactRegistration_DimProject] FOREIGN KEY ([ProjectID]) REFERENCES [dbo].[DimProject] ([ProjectID]),
20    CONSTRAINT [FK_FactRegistration_DimDate] FOREIGN KEY ([DateID]) REFERENCES [dbo].[DimDate] ([DateID]),
21    CONSTRAINT [FK_FactRegistration_DimBoardID] FOREIGN KEY ([DimBoardID]) REFERENCES [dbo].[DimBoard] ([DimBoardID])
22 );

```

Figuur 16, SQL-query aanmaken tabellen en relaties in Visual Studio

Gegevens wegschrijven naar Azure SQL Database

Om vervolgens alle data weg te schrijven naar de database, heeft Marcel van der Linden, medewerker van Strukton Rolling Stock, een PHP-script gemaakt waarmee de .XML-bestanden naar de SQL-database weggeschreven kunnen worden.

Aan de hand van het PHP-script zijn verschillende tests uitgevoerd met .XML-bestanden van verschillende omvang. Dit heeft geleid tot verschillende doorontwikkelingen op detail- en kleine aanpassingen in het datamodel. Het volledige PHP-script is terug te vinden in bijlage XIV.

Na de laatste aanpassingen is het PHP-script op de backoffice van Strukton onafgebroken uitgevoerd waarbij alle beschikbare data geïmporteerd is in de Azure SQL Database. Doordat het importeren van de data een op een virtuele machine is uitgevoerd, heeft dit geen effect gehad op overige werkzaamheden van zowel Strukton als voor mij. Om het importeren van de data te versnellen, heb ik gekozen om het aantal Data Throughput Units (DTU's) van de Azure

SQL Database op te schalen. Voor een kleine verhoging van de resourcekosten, is op deze wijze de verwerkingssnelheid van de database verhoogd.

Na het importeren van een kleine hoeveelheid bestanden, ben ik alvast begonnen met de ontwikkeling van het dashboard in Power BI op basis van de op dat moment beschikbare data. De resultaten die hieruit voort zijn gekomen en de wijze waarop worden toegelicht in hoofdstuk 11.

10.4 Conclusie

Op basis van de bevindingen met de solution template valt te concluderen dat de elementen uit de Cortana Intelligence Suite nog niet volledig geschikt zijn voor iedere situatie. Er zit een grote afhankelijkheid in het format van de data wat gehanteerd wordt. Een conversie tool kan hierbij de oplossing bieden, maar blijkt in de situatie van Strukton met de proof-of-concept niet het gewenste resultaat op te leveren.

In situaties waarbij de organisatie, al dan niet in de industriële sector, nog geen format voor de data heeft, of nog geen historische data heeft opgebouwd, kan er op basis van de eisen van de Cortana Intelligence Suite aangestuurd worden op .CSV of .JSON-bestanden.

Wanneer de toepassing van de Cortana Intelligence Suite zich meer richt op het verkrijgen van inzichten over grote hoeveelheden data, kunnen we spreken van business intelligence en analytics toepassingen. In dit scenario is er een gelijkwaardige toepassing als waar in de proof-of-concept sprake van is. Hierbij wordt enkel de Azure SQL Database uit de Cortana Intelligence Suite gebruikt voor het centraal opslaan en bij elkaar brengen van de data. Aansluitend hierop wordt er met Power BI verbinding gemaakt met de database om de gegevens om te zetten in relevante informatie.

Kortom, toepassing van de volledige Cortana Intelligence Suite brengt de nodige eisen met zich mee. Deze eisen zijn op het gebied van inrichting en dataformat. De situatie van de proof-of-concept voldeed niet aan deze eisen, en de ingeslagen weg met de work-around bood hierbij geen passende oplossing.

De aansluiting van de gekozen oplossing bij de behoeften van Strukton Rolling Stock wordt beschreven in hoofdstuk 11.

11 Werkplaats: Visualisatie in Power BI

In de Cortana Intelligence Suite is Microsoft Power BI de laatste stap in de aaneenschakeling van Azure diensten. Power BI zorgt hierbij voor de visuele weergave van de gegevens. Dit is te zien in de schematische weergave zoals getoond in Figuur 15.

De stap om van gegevens relevante informatie te maken, worden in Power BI relevante waarden tegen elkaar uit te gezet in bijvoorbeeld een staaf-, cirkel- of lijndiagram. Power BI biedt echter nog veel meer functies. In dit hoofdstuk wordt ingegaan op de mogelijkheden die Power BI biedt in combinatie met de betreffende set aan gegevens van Strukton Rolling Stock.

In hoofdstuk 11.1 wordt allereerst ingegaan op de gebruikte methode en werkwijze om de gegevens om te zetten in relevante informatie met Power BI. Hierop volgend worden de visualisatiekeuze voor de Power BI Visualisaties in Power BI toegelicht in hoofdstuk 11.2. Aansluitend hierop worden het demo dashboard en het operationele dashboard getoond in hoofdstuk 11.3. Tot slot bevat hoofdstuk 11.4 een conclusie met betrekking tot de visualisaties in Power BI voor de proof-of-concept.

11.1 Methode en werkwijze

Gedurende voorgaande opdrachten bij Centric is al kennis opgedaan van Microsoft Power BI en de mogelijkheden. De release snelheid van Power BI is in verhouding tot andere BI gerelateerde producten en diensten zeer hoog. Microsoft brengt maandelijks een nieuwe versie uit van Power BI waarbij nieuwe mogelijkheden en visualisaties toegevoegd worden. Doordat Microsoft bij de doorontwikkeling van Power BI gebruik maakt van een online community, is de mogelijkheid tot inbreng van nieuwe ideeën hoog.

Buiten de standaard aangeboden visualisaties, maakt Power BI het ook mogelijk om andere visualisaties van externe ontwikkelaars te implementeren in het dashboard. Met deze extra visualisaties worden de mogelijkheden van Power BI enorm uitgebreid.

Tijdens een overleg met Han Geerligts en Marcel van der Linden is afgesproken om op experimenterende wijze een dashboard in Power BI te ontwikkelen. Deze keuze is gemaakt om het einddoel waar naar toegewerkt wordt zichtbaar te maken voor alle relevante stakeholders. Vanuit de experimenterende ontwikkeling is het demo dashboard ontstaan.

Met het demo dashboard is binnen Strukton Rolling Stock meer bekendheid rondom het afstudeerproject gecreëerd en direct feedback gevraagd op basis van de getoonde visualisaties. Hierbij zijn aanvullende uitbreidingen op de data genoemd, maar ook visualisaties op basis van de huidige dataset.

Samenvattend wordt deze werkwijze trial-and-error ontwikkeling genoemd. Deze methode kan omschreven worden als uitproberen en daarbij leren van fouten. Het demo dashboard is op basis van eigen inschattingen van de data ontwikkeld, waarna deze is aangepast op basis van de feedback. Deze wijze van ontwikkeling is enkel van toepassing geweest op de informatiebehoefte en niet op het gebruik van Power BI. Kennis van Power BI was reeds aanwezig wat bij de demo gemakkelijk toe te passen was. (Wikipedia, 2016)

De feedback van de gebruikers heeft uiteindelijk gezorgd voor input en veranderingen in de demo die wederom bij de gebruikers aangeboden is voor feedback. Na beschikbaarheid van de volledige dataset, is deze informatiebehoefte uit het dashboard omgezet in een dashboard met de volledige dataset.

11.2 Visualisaties in Power BI

Voor zowel de demo als de operationele omgeving is een dashboard ontwikkeld in Power BI. Deze tooling biedt veel verschillende visualisaties binnen de standaard mogelijkheden, maar daarbuiten ook nog custom visualisaties die van de Power BI website te downloaden zijn. Echter vond ik het lastig om de juiste grafiek te kiezen bij de beschikbare data. Om hier meer duidelijkheid over te krijgen heb ik verschillende bronnen gebruikt die meer informatie geven over een duidelijk, overzichtelijk en veel gebruikt type grafiek of andere visualisatie.

In een document opgesteld door dhr. Welman wordt het volgende gezegd over grafieken:

“Grafieken zijn een belangrijk hulpmiddel bij het analyseren van cijfers. Grafieken kunnen informatie helder en duidelijk weergeven en de kracht ervan moet niet onderschat worden. Trends en overzichten kunnen in grafieken weergegeven worden en afwijkende waarden vallen er snel in op.”

(Welman, n.d.)

In het citaat van dhr. Welman wordt exact benadrukt waar het in het dashboard van de PoC om draait, namelijk het analyseren van cijfers om hiermee trends te kunnen ontdekken waarbij afwijkende waarden direct opvallen. In dit document wordt voornamelijk gefocust op de grafieken in Microsoft Excel. Aangezien Power BI kan worden gezien als doorontwikkeling van Power Pivot, Query en View uit Excel, zijn hier veel overeenkomsten in te vinden.

Naast het document van dhr. Welman, is een onlinebron gebruikt voor de keuze van visualisaties. Deze website adviseert grafiektypen op basis van type informatie. De volgende typen worden hierbij onderkent:

- Trend;
- Relatie;
- Vergelijking;
- Distributie;
- Compositie.

(Juice Analytics, n.d.)

11.3 Power BI dashboards

Demo dashboard

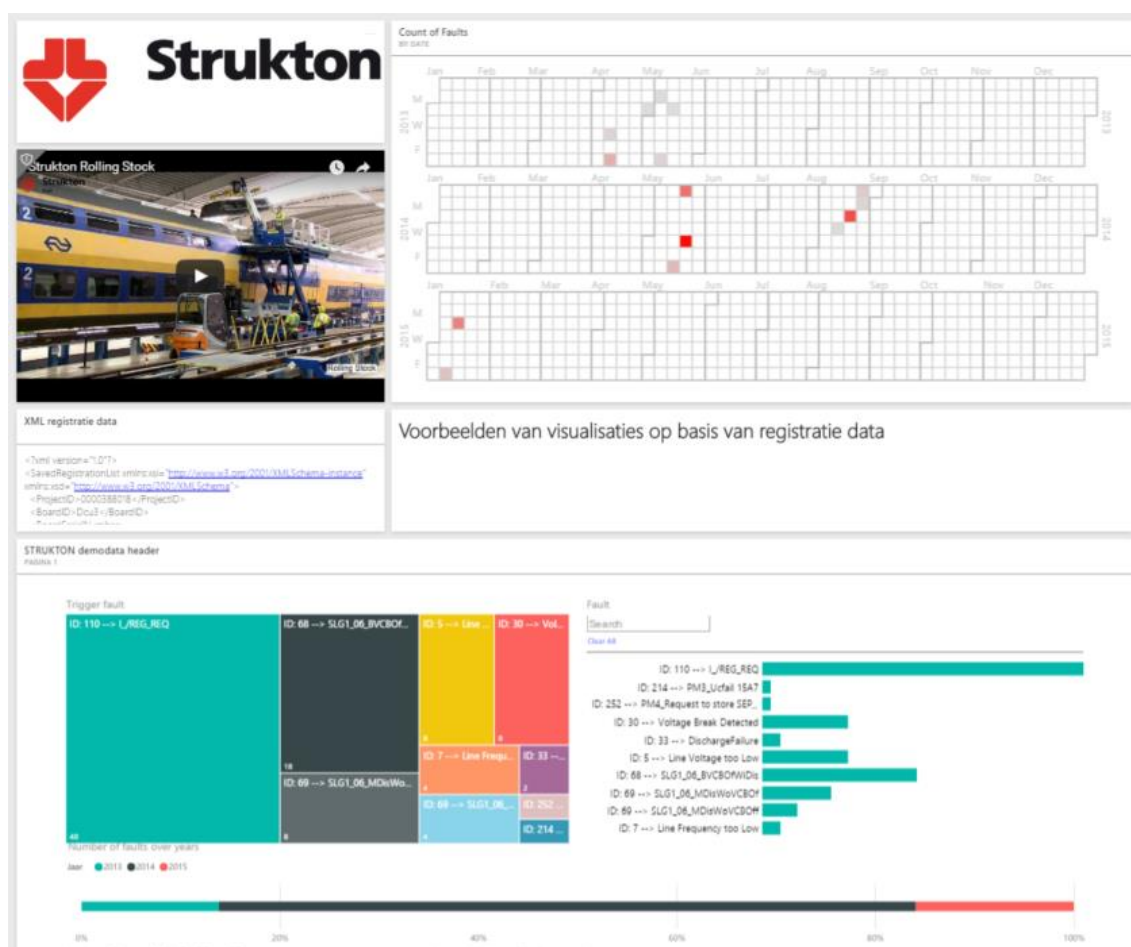
In het demo dashboard, wat naar eigen inzicht gemaakt is, is enkel de header van de registraties gebruikt. De header van de registraties omvat alle data uit de registratie met uitzondering van de samples, ook wel signalen zowel analoog als digitaal. Om een representatieve demo te kunnen tonen zonder vooraf de database en de import van alle data op orde te hebben, is handmatig een kleine set met registraties samengesteld. Al deze registraties zijn in één tabel op de Azure SQL Database geïmporteerd waar Power BI verbinding mee maakt.

Met de demo is direct getest wat de mogelijkheden zijn met betrekking tot het verbinden met een externe database vanuit Power BI. In de standaard situatie kiest Power BI ervoor om alle data te importeren in het Power BI-bestand. Uit ervaring is gebleken dat de gecreëerde bestanden hierdoor een enorme omvang kunnen aannemen, en mogelijk zelfs onleesbaarheid van het bestand tot gevolg heeft. Bij de proof-of-concept is om die reden bewust gekozen voor een direct-query.

Bij gebruik van een direct-query bevat het bestand enkel de verbindingsgegevens, query's en visualisaties, maar geen data. Met de verbindingsgegevens wordt bij het openen van het bestand verbinding gemaakt met de onlinedatabron om daar de query's uit te voeren. In het geval van de proof-of-concept is de onlinedatabron de Azure SQL Database.

Het demo dashboard is tijdens een maandelijks terugkerend overleg bij Strukton gepresenteerd aan de gebruikersgroep. Tijdens dit overleg is uitleg gegeven bij de visualisaties op basis van de data, en de wijze hoe ik hier toe ben gekomen. De feedback die hierop ontvangen is, heeft de input gevormd voor de doorontwikkeling in het operationele dashboard.

In het demo dashboard is gebruik gemaakt van meerdere reports. Onderstaand figuur toont een deel van het demo dashboard Het volledige demo dashboard is terug te vinden in bijlage XV.



Figuur 17, Demo dashboard Strukton Rolling Stock

Operationeel dashboard

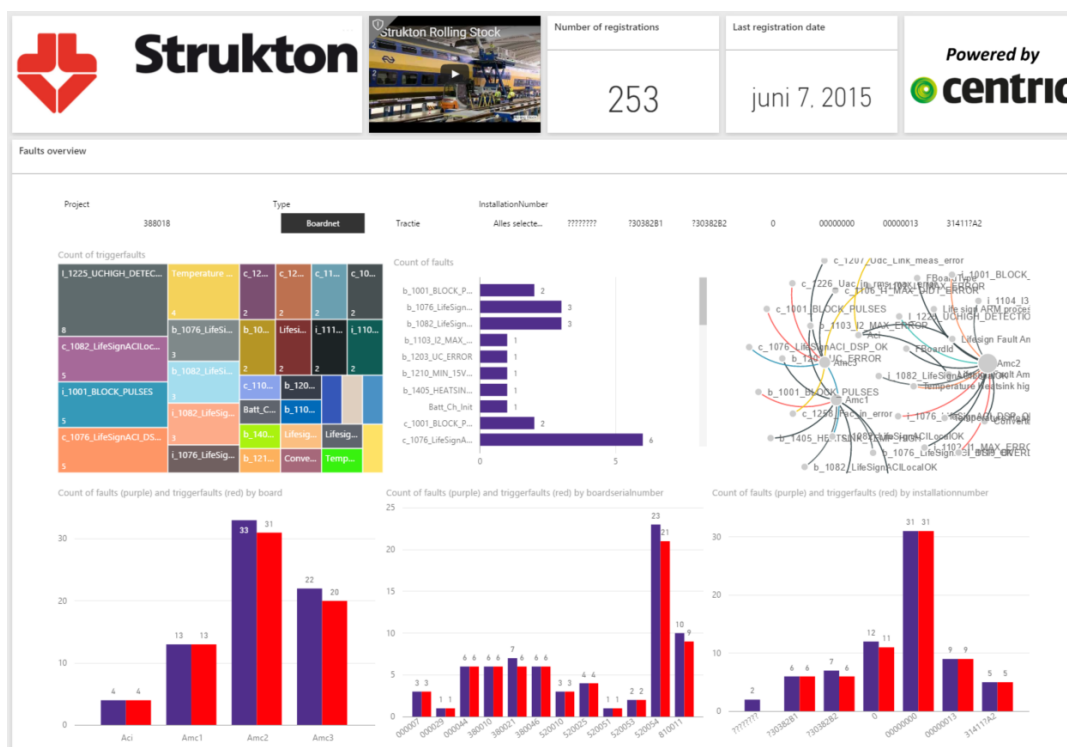
Het eindproduct van de proof-of-concept en het dashboard waarin de volledige dataset geladen is, vormt het operationele dashboard. In dit dashboard is de feedback van de gebruikersgroep verwerkt tot de huidige versie van het dashboard.

Het operationele dashboard maakt gebruik van direct query waarbij een directe verbinding gemaakt is met de Azure SQL-database. Het dashboard maakt bij iedere keer dat het opgevraagd wordt opnieuw verbinding met de database. Hierdoor heeft de eindgebruiker continu inzicht in de meest actuele versie van het dashboard. Hiermee wordt het ook mogelijk gemaakt dat de nieuwe registraties in de toekomst automatisch vanuit de Strukton backoffice worden toegevoegd aan de database en verschijnen in het dashboard.

Enkele verbeterpunten in de operationele versie ten opzichte van de demoversie zijn:

- Lijndiagrammen vervangen door staafdiagrammen omdat gaten in de data met een lijndiagram niet zichtbaar zijn;
- Fouten per installatienummer (locomotief) inzichtelijk maken;
- Fouten per softwareversie;
- iButtonnummers tonen bij fouten. Deze waarde bevat informatie over de status van de machine;
- Onderscheid en filtermogelijkheden in tracties en boardnetten;
- Filtermogelijkheden op installatienummer;
- Filtermogelijkheden op project met het oog op uitbreiding van de dataset met andere projecten van andere klanten van Strukton Rolling Stock.

In het dashboard is gebruik gemaakt van meerdere reports. Onderstaand figuur toont een deel van het dashboard. Het volledige dashboard is terug te vinden in bijlage XVI.



Figuur 18, Operationeel dashboard registratie data Strukton Rolling Stock

11.4 Conclusie

De operationele versie van het dashboard is wederom gedeeld met de gebruikersgroep waarna om feedback gevraagd is. Het dashboard is door de gebruikersgroep positief ontvangen. Hierbij werd door verschillende gebruikers aangegeven dat zij hiermee inzichten krijgen over grotere hoeveelheden data, waar zij voorheen niet over beschikten.

In de reacties van de gebruikers zijn nog enkele verbeterpunten en gewenste uitbreidingen genoemd die op dit moment niet in het dashboard aanwezig zijn. Hierbij kan gedacht worden aan data uit andere projecten en het toevoegen van locatie gegevens van de registraties. Voor de PoC zijn deze punten buiten de scope geplaatst in verband met de beperkte resterende tijd.

Andere feedback punten van gebruikers, waar het ging om de weergave van een visualisatie of de daarbij behorende filtermogelijkheden zijn direct naar wens aangepast in de operationele versie. Deze wijzigingen zijn in de schermafbeeldingen niet zichtbaar maar enkel in de interactieve weergave voor de gebruikers.

Het onderdeel wat in de huidige dashboards nog niets mee gedaan is, maar wel in de database aanwezig is, betreffen de meetwaarden die behoren bij iedere registratie. Deze meetwaarden geven voor medewerkers van Strukton informatie over de oorzaak van een fout. Echter is geconcludeerd dat deze data waardevol is voor toepassing van machine learning en nietszeggend is om zonder bevraging of daaraan verbonden conclusies te visualiseren in het dashboard. Deze data is wel meegenomen bij het importeren van de data in de database voor toekomstige doeleinden.

12 Lab: Aansluiting Cortana Intelligence Suite bij de behoeften van Strukton Rolling Stock

Met de proof-of-concept is voor Strukton Rolling Stock een werkende situatie geleverd waarin van alle registraties een overzicht wordt gegeven. Maar wat heeft dit nu precies voor de gebruikers opgeleverd? En wat kunnen ze hier mee?

In dit hoofdstuk wordt toegelicht wat de toegevoegde waarde van de uitgevoerde PoC is voor de gebruikers van Strukton Rolling Stock. Hiervoor wordt beschreven op welkewijze deze informatie verzameld is in hoofdstuk 12.1. Aansluitend hierop wordt in hoofdstuk 12.2 toegelicht welke toegevoegde waarde met de huidige situatie van de PoC is bereikt voor Strukton. Vervolgens beschrijft hoofdstuk 12.3 de mogelijke uitbreidingen op de PoC waar Strukton interesse in heeft. Tot slot wordt op basis hiervan een conclusie gevormd in hoofdstuk 12.4.

12.1 Werkwijze

Om meer informatie te verkrijgen over de toegevoegde waarde van de PoC voor de gebruikersgroep, is een inventarisatie gedaan bij deze groep. Hierbij is gevraagd het operationele dashboard te bekijken en op basis hiervan de toegevoegde waarde voor hun dagelijkse werkzaamheden te omschrijven. Daarnaast is gevraagd om eventuele aanvullingen voor toekomstige doorontwikkelingen.

Alle verzamelde punten zijn meegenomen in een overleg met Strukton en Centric. In dit overleg is besproken wat er bereikt is gedurende de uitvoering van de PoC, wat dit opgeleverd heeft, uitbreidingen en een eventueel vervolgtraject.

12.2 Toegevoegde waarde van bereikte resultaat

Het operationele dashboard is door de gebruikersgroep erg positief ontvangen. Zo krijgen zij nieuwe inzichten over een groter tijdsbestek die zij voorheen niet hadden. De feedback betrof enkele kleine aanpassingen zoals beschreven in hoofdstuk 11 en zijn direct aangepast.

Andere toevoegingen die vanuit de gebruikersgroep gegeven werden, hadden betrekking op het voorspellen van meetwaarden en fouten. Hierbij kon worden aangegeven dat dit in een eerder stadium, in overleg met aanspreekpunt bij Strukton, buiten de scope geplaatst is.

In eerste instantie was het niet voor alle gebruikers mogelijk om het dashboard te openen. Doordat het dashboard vanuit de Centric tenant gedeeld is met de medewerkers van Strukton in hun eigen tenant, staat Microsoft het niet toe dat deze gebruikers alle rechten hebben. Ook zijn er elementen in het dashboard aanwezig waarvoor een Power BI-Pro licentie vereist is.

Na kleine aanpassingen hebben de gebruikers het dashboard kunnen bekijken en hun feedback kunnen geven. Alle ontvangen feedback was positief gestemd. René van den Hoek, technology manager bij Strukton Rolling Stock, heeft dit het meest concreet onder woorden kunnen brengen.

“We kunnen over een grote hoeveelheid data nieuwe inzichten verkrijgen die wij in het verleden niet hadden of lastig konden krijgen. De verwachtingen met het verder aanvullen van de data zijn positief. De resultaten van steekproeven uit het verleden zien wij in het dashboard terug.

...

Met het dashboard krijgen wij een overzicht van de fouten om deze systematischer te benaderen. We hopen in de toekomst met het verder aanvullen van de database nieuwe inzichten te krijgen en patronen te herkennen waar we nog nooit aan gedacht hebben.”

Om dergelijke problemen voor de toekomst op te lossen, wordt de database en het dashboard op korte termijn gemigreerd naar de Strukton tenant. Hiermee wordt het voor de (eind)gebruikers gemakkelijker om het dashboard te benaderen.

12.3 Uitbreidingen

In de huidige versie van de proof-of-concept zit (nog) geen voorspellende waarde. Ook is alle getoonde informatie nu (nog) afhankelijk van de interpretatie van de gebruiker. René van den Hoek heeft zijn verwachtingen geuit, waarin een uitbreiding wordt gerealiseerd met Azure Machine Learning. Met toepassing van machine learning, kan met de juiste algoritmen de data geïnterpreteerd worden en op basis daarvan een conclusie vormen die aan de (eind)gebruiker gepresenteerd wordt in het dashboard. Toepassing van machine learning is door René van den Hoek en Han Geerligts aangedragen als volgende stap voor mogelijk een nieuwe afstudeer stagiair bij Centric.

Bij alle registraties wordt al een locatie doorgestuurd in een apart bestand. Een uitbreiding op het huidige dashboard is het toevoegen van de locatiegegevens. Door de locatie van de fouten op een kaart te plotten, kan inzichtelijk worden gemaakt of er in een bepaald gebied veel fouten zijn. Dit kan leiden tot externe oorzaken van de fouten waar in het verleden nog geen informatie over was. Hierbij kan gedacht worden aan problemen met de spoorweginfrastructuur of weersomstandigheden in bepaalde gebieden. Dit leidt ook direct tot de volgende mogelijke uitbreiding.

Wanneer er locatiegegevens beschikbaar zijn en aan de registraties gekoppeld zijn, kunnen er meteorologische gegevens van de betreffende locaties aan toegevoegd worden. Om waardevolle meteorologische informatie te tonen en de verbanden hiertussen te laten zien, is de beschikbaarheid van locatiegegevens vereist. Strukton beschikt niet over meteorologische data, echter is dit veelal open-data waardoor dit alsnog vanuit een onlinebron aan de dataset te koppelen is. De beschikbaarheid van deze data voor de gebieden waar de klanten van Strukton zich bevinden en de nauwkeurigheid hiervan vereisen nader onderzoek.

12.4 Conclusie

Op basis van de reacties van Strukton kan geconcludeerd worden dat de inzichten die geboden worden met het Power BI-dashboard een welkome aanvulling is op de huidige informatievoorziening. De complexe data-analyse die voorheen steekproefsgewijs uitgevoerd werden, zijn met de inzichten in dit dashboard overbodig geworden. De database en het dashboard worden gezien als basis fundering vanwaaruit doorontwikkeld en uitgebreid kan

worden. Hierbij is een uitbreiding met locatiegegevens een eerste stap ter uitbreiding van de informatievoorziening.

De toepassing van machine learning is een complexer geheel wat om een onderzoekend deel vraagt. Strukton ziet dit graag wederom in de vorm van een afstudeer stagiair bij Centric. Hierbij wordt de kennis opgebouwd binnen Centric, en profiteert Strukton van de resultaten die daaruit voortvloeien.

In onderstaande tabel met de MoSCoW prioritering van de informatiebehoeften zoals vastgesteld in hoofdstuk 9.2. Hierin is zwarte tekst de ingevulde en rode de niet ingevulde informatiebehoefte.

	Must-have:	Should-have:	Could-have:	Would-have:
1	Welke fout treed er het meeste op?	Is er een tijdsrange of periode waarin de meeste fouten optreden.	Locatiegegevens van de registraties om mogelijke oorzaken te achterhalen van de fouten.	Verkoopgegevens om de verhouding tussen de vloot en de fouten inzichtelijk te maken.
2	Welke trigger leidt tot de meeste registraties, en hoeveel zijn dat er?	Patroonherkenning en fouten aan zien komen op basis van de signalen bijbehorend bij de registraties.		Onderhoudsgegevens om te kijken wanneer welke locomotief onderhoud heeft gehad en de relatie te leggen met de fouten.
3	Hoeveel fouten zijn er per board?			Meteorologische data om verbanden en de afhankelijkheid van weersomstandigheden met de fouten inzichtelijk te maken.
4	Hoeveel fouten zijn er per locomotief (installatienummer)?			
5	Welke versie of serienummer van het board vertoont het meeste fouten?			

Tabel 4, Review op MoSCoW indeling informatiebehoeften

13 Showroom: Aansluiting Cortana Intelligence Suite voor de industriële sector

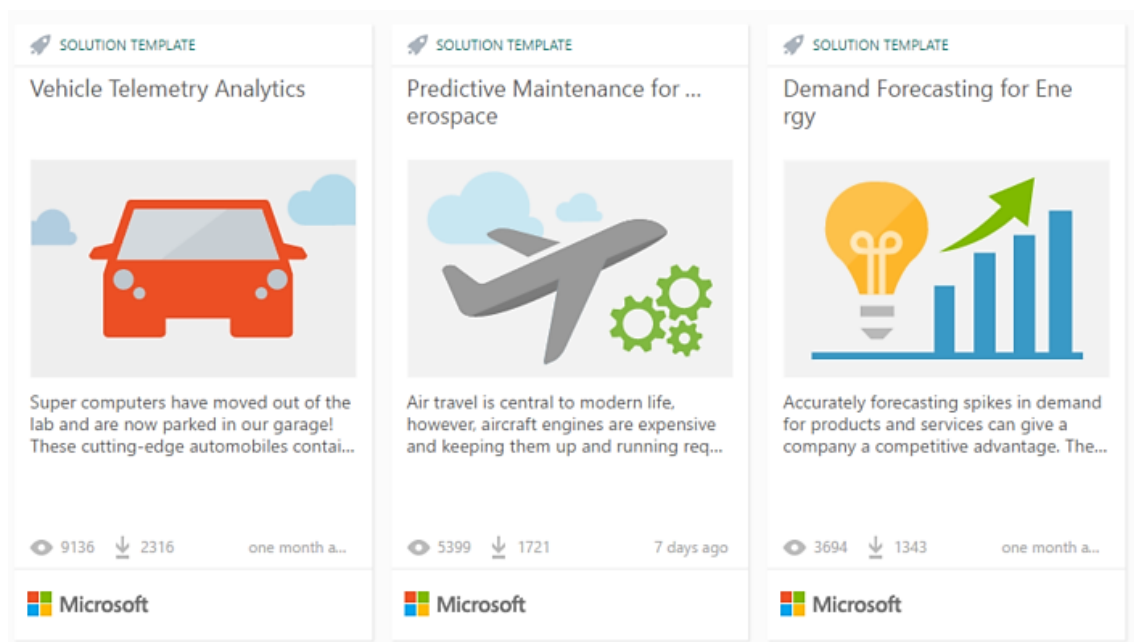
In dit hoofdstuk wordt er gekeken of de geboden toepassing uit de proof-of-concept ook geschikt is voor andere organisaties. Hoofdstuk 13.1 betreft een toelichting op de mogelijkheden binnen de industriële sector zoals beschreven in de onderzoeksvraag. Tot slot wordt hier in hoofdstuk 13.2 een conclusie bij gegeven.

13.1 Toepassing in de industriële sector

De onderdelen uit de Cortana Intelligence Suite zoals deze zijn gebruikt in de proof-of-concept voor Strukton R.S., zijn ook toepasbaar in andere situaties. Hierbij is een toepassing zowel binnen als buiten de industriële sector denkbaar. De gestelde onderzoeksvraag richt zicht echter volledig op de industriële sector.

Solution templates

Er zijn verschillende toepassingen denkbaar waarbij de Cortana Intelligence Suite een oplossing kan bieden. Microsoft biedt vanuit de Cortana Intelligence Gallery, zoals toegelicht in hoofdstuk 10.2, verschillende solution templates waarbij de volledige suite aan Azure cloud diensten wordt gebruikt. De solution templates die worden aangeboden, bevinden zich allen binnen de industriële sector zoals in Figuur 19 te zien is.



Figuur 19, Solution templates uit de Cortana Intelligence Gallery

Toename vraag

De verwachting is dat het aantal toepassingen en de vraag naar analytics toepassingen groeit. Deze verwachting is gebaseerd op fase waarin internet-of-things zich bevindt in de Gartner hype cycle zoals te zien is in Figuur 4 in hoofdstuk 3.4. De huidige fase, peak of inflated

expectations, kenmerkt zich dat het aantal toepassingen en de ervaring groeit waarmee de vraag naar toepassingen met de betreffende technologie toeneemt.

De vraag naar de toepassingen is binnen Centric al merkbaar. Vanuit Centric zijn er al diverse leads waarbij met Power BI een oplossing wordt voorgesteld. Om Power BI van data te voorzien en de inrichting hiervan, is in alle gevallen een toepassing van de Cortana Intelligence Suite mogelijk en denkbaar. Enkele leads bevinden zich in de industriële sector. Daarnaast is er ook een lead in de agrarische sector.

13.2 Conclusie

Er valt te verwachten dat de vraag naar data-analyse en internet-of-things oplossingen gaat toenemen. Deze toename betreft zowel de industriële als de overige sectoren. Vele organisaties beschikken over grote hoeveelheden data waarmee diverse bedrijfsprocessen geoptimaliseerd kunnen worden.

Bij toepassing van de Cortana Intelligence Suite, of elementen hiervan, voor andere bedrijven en in andere sectoren, is beschikking over of de opbouw van domeinkennis zeer belangrijk voor juiste interpretatie van de gegevens.

Ook speelt het dataformat een rol om te bepalen welke onderdelen uit de Cortana Intelligence Suite bruikbaar zijn voor de casus van de organisatie. Op basis van de huidige fase in de Gartner hype cycle is te verwachten dat er diverse doorontwikkelingen plaats gaan vinden. Een verrijking van de ondersteunde dataformats is daarbij niet ondenkbaar.

Daarnaast is de communicatie richting van de data van cruciaal belang voor de toepasbaarheid van de Cortana Intelligence Suite. De suite accepteert enkel inkomende data en maakt het niet mogelijk om op basis van de data acties te pushen naar de IoT devices. Wanneer deze tweerichtingen data communicatie gewenst is, dient er een andere suite van Microsoft gebruikt te worden.

14 Conclusie en aanbeveling: Cortana Intelligence Suite voor Centric

In dit hoofdstuk wordt voor Centric een conclusie gevormd over de toepassing van de Cortana Intelligence Suite. Hiervoor worden allereerst in hoofdstuk 14.1 de doorontwikkelingen besproken waar Centric een rol in kan spelen. Aansluitend daarop wordt standaardisatie van oplossingen besproken in hoofdstuk 14.2. Vervolgens wordt de aanbiedingsvorm die Centric kan bieden aan haar klanten in hoofdstuk 14.3 besproken. Tot slot wordt mijn visie op onderhoud en beheer van dergelijke oplossingen uitgewerkt in hoofdstuk 14.4. Hoofdstuk 14.5 vat de conclusie en aanbevelingen kort samen.

14.1 Doorontwikkeling

Voor Centric is het van belang om te investeren in de ontwikkeling van kennis en kunde op het gebied van analytics. Zoals in hoofdstuk 13.1 is beschreven, neemt de vraag toe. Om aan deze vraag te kunnen voldoen en daarmee ook aan de verwachtingen van de prospects en klanten te kunnen voldoen, is het van belang dat deze kennis opgebouwd wordt.

Centric moet zich hierbij niet alleen focussen op de groei van analytics in de industriële markt. Ook in de retail en andere markten neemt de vraag toe. Voor iedere branche is een toepassing van analytics denkbaar. De ontwikkeling moet zich dus niet alleen focussen op de technische kennis van analytics, maar ook domeinkennis van de branche. Deze domeinkennis is van belang om de klantvraag goed te kunnen begrijpen en te vertalen naar een passende oplossingen voor de klant.

14.2 Standaardoplossingen

Het ontwikkelen van standaardoplossingen voor iedere branche kan een wijze zijn waarop Centric de analytics propositie kan vormgeven. Maar zoals gebleken is uit de proof-of-concept zijn de mogelijkheden voor standaardoplossingen beperkt. Door de afhankelijkheid van het bestandsformat en het type gegevens die de klant heeft, kan een standaardoplossing voldoen of falen.

Iedere klantvraag is uniek, wat de mogelijkheden voor standaardoplossingen beperkt. Aangezien de Power BI-visualisaties uitgaan van de kolomnamen uit de database, is een standaard dashboard als aanvulling op maatwerk software of specifiek vraagstuk niet mogelijk. Bij slechts geringe variatie in kolomnamen zullen visualisaties in Power BI niet meer werken of onjuiste informatie tonen.

Er zijn wel andere situaties denkbaar waarin standaardisatie mogelijk is. Een voorbeeld hiervan is een analytics als aanvulling op een bestaand standaardproduct. Bijvoorbeeld een uitbreiding op Microsoft Dynamics CRM. In de situatie van Dynamics CRM zijn de database namen veelal gelijk aan elkaar en kan een standaardoplossing relatief eenvoudig ingericht worden.

Zeker in de beginfase waarin Centric zich nog sterk ontwikkelt op het gebied van analytics, betreft de meerderheid van de analytics oplossingen maatwerk. Mogelijk ontwikkelt de Cortana

Intelligence Suite en de kennis binnen Centric zich zodanig, dat standaardisatie in de toekomst wel mogelijk is.

14.3 Aanbiedingsvorm en inrichting

Centric breed wordt er steeds meer gestreefd naar het aanbieden van de producten in een SaaS vorm, software as a service. Hierbij betaald de klant periodiek een bedrag en wordt daarbij niet belast met beheer, onderhoud of een hoge aanschafprijs. SaaS-oplossingen zijn veelal cloud oplossingen maar daarbij houdt het niet op. SaaS is namelijk meer dan alleen software in de cloud. Het vraagt om een andere manier van denken waarbij standaardisering het sleutelwoord is.

Het visiedocument van Centric ten aanzien van SaaS is bijgevoegd in bijlage XVII.

Echter is het aanbieden van analytics vanwege de uiteenlopende informatiebehoeften erg lastig. Door deze uiteenlopende behoeften die klant specifiek zijn, spreken we van maatwerk. Bij de ontwikkeling van software worden generieke componenten ingezet die de snelheid van de ontwikkeling vergroten, maar ook leiden tot hogere kwaliteit. Dezelfde werkwijze kan ook worden toegepast bij het aanbieden van analytics. Gebruik van standaard elementen bij analytics kunnen bestaan uit volledig ingerichte databases waarin standaardelementen zoals datum en tijd tabellen al aanwezig zijn.

Multitenancy, waarbij één centrale omgeving (cloud) gedeeld wordt met meerdere organisaties, is niet bevorderlijk. Multitenancy kan worden ingezet om ontwikkelingskosten te verlagen door hergebruik van elementen en klanten te laten betalen op abonnementsbasis voor het gebruik ervan. Echter kan dit met betrekking tot data-analyse gevoelig liggen, omdat mogelijk concurrentiegevoelige data fysiek op dezelfde omgeving opgeslagen wordt. De besparing in ontwikkelingskosten voor Centric weegt hierbij niet op tegen het risico dat klanten gevoelige data van elkaar onder ogen krijgen en de directe gevolgen hiervan. In de ideale situatie blijft de data binnen de tenant van de klant, ook wel in de cloud, om hiermee het risico op datalekken te voorkomen.

Door onderhoud en beheer aan te bieden bij de klanten, kan er gezorgd worden voor een constante inkomstenstroom voor Centric. Dit kan in de vorm van een servicelevel agreement (SLA) of een strippenkaart. De keuze voor een van beide opties dient bij de klant te liggen. Voor een klant waarbij de dagelijkse bedrijfsvoering van de informatievoorziening afhankelijk is, kiest hierbij eerder voor een servicelevel agreement, waarin wederzijdse verwachtingen vastgelegd worden. Een SLA brengt daarentegen hogere kosten met zich mee dan een strippenkaart waar een X-aantal uren door de klant vooraf gekocht wordt.

14.4 Onderhoud en beheer

Na realisatie van een analytics oplossing voor een klant, dient dit ook onderhouden en beheert te worden. Hierbij kan worden gedacht aan uitbreiding, aanpassing en op- of afschalen van de capaciteit. Dit beheer moet zowel plaats vinden op de database als overige analytics onderdelen zoals het dashboard. Dit kan bestaan uit het aanpassen van rechten voor gebruikers, maar ook het aanpassen en toevoegen van visualisaties op basis van wensen. Hierbij behoort ook het oplossen van problemen wanneer er geen beschikbaarheid is van de informatie die valt onder onderhoud en beheer.

14.5 Samenvatting conclusie en advies

Samengevat kan Centric de propositie rondom de toepassing van de Cortana Intelligence Suite als volgt omschrijven:

Centric wordt geadviseerd om analytics en IoT oplossingen als aanvulling op de aan al aangeboden proposities zoals CRM als standaardproduct aan te bieden. De standaard is hierbij voor de klant een eerste kennismaking met de mogelijkheden van analytics en IoT. Voor uitbreidingen op de standaard of specifieke vraagstukken wordt het aanbieden van maatwerk geadviseerd. Maatwerk is van belang om volledig aan te sluiten op de situatie en behoeften van de klanten.

De aanbiedingsvorm dient hierbij binnen de tenant van de klant te zijn om problemen met data beveiliging te voorkomen en de schijn niet tegen te krijgen. Bij deze aanbiedingsvorm zijn alle licenties en kosten bij de klant belegd en betreft dit geen zorg van Centric.

Aanvullend hierop wordt de continuïteit van de beschikbaarheid gewaarborgd in een service level agreement op basis van de behoeften van de klant, afgezet tegen de capaciteiten van Centric om in deze behoeften te kunnen voldoen. Een service level agreement is de ideale situatie voor Centric omdat dit een constante inkomstenstroom genereert. In situaties waarbij de corebusiness (dagelijkse bedrijfsvoering) van de klant niet afhankelijk is van het geleverde product, kan een strippenkaart de oplossing bieden voor de klant om op kosten te besparen.

15 Toekomstig werk

Gedurende de uitvoering van dit onderzoek, zijn er een aantal zaken uit de Cortana Intelligence Suite buiten scope geplaatst. Deze onderdelen zijn buiten scope geplaatst, omdat hier binnen het gestelde termijn geen tijd voor was, of simpelweg niet mogelijk was om technische redenen. Dit hoofdstuk geeft inzicht in de onderdelen die in dit onderzoek buiten beschouwing zijn gelaten. Bij de genoemde onderdelen wordt kort toegelicht waarom deze buiten beschouwing zijn gelaten. Alle hieronder genoemde onderdelen betreffen elementen uit de Cortana Intelligence Suite.

Machine learning

Dit onderdeel is gedurende de uitvoering van de proof-of-concept buiten scope geplaatst. Door opgelopen vertraging met de data conversie en het wegschrijven van de data in de Azure SQL Database, was er weinig tijd over om machine learning toe te passen op de gemeten signalen uit de registraties.

De toepassing van machine learning had in dit onderzoek als doel om de signalen te analyseren en hierin patronen te herkennen om hiermee voorspellingen te kunnen doen. Overigens is vooraf de toepassing van machine learning met de beperkte dataset in twijfel getrokken. Doordat de gegeven set momentopnamen bevat, kan het zo zijn dat het verloop van de meetsignalen waarin de aanloop naar de fout te zien is, niet in de data is meegenomen.

Strukton ambieert een continue datastroom waarmee de mogelijkheden vergroot worden.

Perceptuele intelligentie

Dit onderdeel uit de Cortana Intelligence Suite maakt het mogelijk om gezichtsuitdrukkingen, spraak en tekst te analyseren. Dit is tijdens de proof-of-concept niet meegenomen, omdat deze gegevens niet aanwezig zijn in de casus van Strukton Rolling Stock en niet relevant zijn geacht.

Live dashboards met streaming analytics

Met het element Azure Stream Analytics wordt het mogelijk gemaakt om data live weer te geven in een Power BI-dashboard. De plaats van Azure Stream Analytics is te zien in Figuur 12. Dit onderdeel is niet gebruikt omdat de dataset van Strukton een datadump uit het verleden bevat, en er geen continue datastroom aan nieuwe gegevens binnen komt via de Azure Event Hub.

Ongestructureerde data met Azure HDInsights (Hadoop)

Via Azure HDInsights (Hadoop) wordt het mogelijk gemaakt om ongestructureerde data te analyseren. Aangezien alle data van Strukton gestructureerd van aard is, en alles in hetzelfde format opgeslagen zijn, was er geen sprake van ongestructureerde data en geen aanleiding om Azure HDInsights binnen de proof-of-concept te gebruiken.

Persoonlijke digitale assistent, Cortana

Cortana maakt het mogelijk om in spreektaal vragen te stellen over de data. Hiervoor is echter wel een apparaat met Windows 10 of Windows mobile als operating system nodig. De volledige kantoorautomatisering van Strukton gebruikt Windows 7 als operating system. Hierdoor is het niet relevant geacht om dit onderdeel te testen in de proof-of-concept. Strukton is bezig met een migratie van haar kantoorautomatisering naar Windows 10 waarmee Cortana mogelijk op korte termijn een aanvulling kan bieden.

Literatuurlijst

- About Tech - Mike Chapple, database expert. (2014, december 16). *About Tech - Databases - Facts vs. Dimensions - Understanding Key Business Intelligence Terms*. Opgehaald van About Tech: <http://databases.about.com/od/datamining/a/Facts-Vs-Dimensions.htm>
- Alex Feijt - Research Company. (2013, december 18). *Basiskennis onderzoek: Kwalitatief onderzoek, wat is het en waarom doen we het?* Opgehaald van Research Company: <http://www.rcompany.nl/basiskennis-onderzoek-kwantitatief-onderzoek-wat-is-het-en-waarom-doen-we-het/>
- Alles over marktonderzoek. (sd). *Alles over marktonderzoek - Diepte interview*. Opgehaald van <http://www.allesovermarktonderzoek.nl/onderzoeksmethoden/diepte-interview>
- Anne Coppens, Marjolein Jacobs, Tineke Jacobs, Ralph Niels, Nynke Verhoeven. (2015, april). *Proeven van onderzoek - De methodenkaart in de beroepspraktijk van ICT en Media*. Zevenaar: Informatica en communicatie academie (ICA), Hogeschool van Arnhem en Nijmegen.
- Belastingdienst. (sd). *Documenten Belastingdienst*. Opgehaald van Lijst bedrijfsactiviteiten/NACE-codes: http://download.belastingdienst.nl/belastingdienst/docs/bedrijfsactiviteiten_nace_codes_ob4111b4pl.pdf
- Centraal Bureau voor de Statistiek. (2016). *Begrippen overzicht - Innovatie*. Opgehaald van Centraal Bureau voor de Statistiek - methoden en technieken: <http://www.cbs.nl/nl-NL/menu/methoden/begrippen/default.htm?conceptid=2703>
- Centraal Bureau voor de Statistiek. (2016). *Methoden en begrippen*. Opgehaald van Centraal Bureau voor de Statistiek: <http://www.cbs.nl/nl-NL/menu/methoden/begrippen/default.htm>
- Centric - Dick van Straaten. (2016, april 15). *Blog: SaaS, een relatie in de cloud*. Opgehaald van Centric - Craft - Blogs: <http://www.centric.eu/NL/Default/Craft/Blogs/2016/04/15/SaaS-een-relatie-in-de-cloud>
- Centric - Prof. Dr. Ben van Lier. (sd). *Ben van Lier, expert Internet of Things*. Opgehaald van Centric.eu - Experts: <http://www.centric.eu/NL/Default/Themas/Experts/Ben-van-Lier>
- Centric Netherlands B.V. - Floris van Dijk. (2015). *Microsoft Dynamics CRM: Verbeter uw klantrelaties*. Opgehaald van Centric.eu: <http://www.centric.eu/NL/Default/Software-Solutions/Standaardsoftware/Microsoft-Dynamics-CRM>
- Centric Netherlands B.V. - Hans Hengst - Partner Manager. (sd). *Partnership Microsoft en Centric*. Opgehaald van Centric.eu: <http://www.centric.eu/NL/Default/IT-Outsourcing/Partners/Partnerverhaal?id=6396eb41-745c-4e19-8bc0-29749f42d6a2>
- Centric Netherlands B.V. (2015). *Over Centric*. (M. Aanhaanen, Redacteur) Opgeroepen op september 16, 2015, van Centric.eu: <http://www.centric.eu/NL/Default/Over-Centric>
- Centric Netherlands, Solution Engineering, Will Loermans. (2015). *Rapport Centric SE Machine Analysis propositie Industrie*. Capelle aan den IJssel.
- Data Mining 4U. (sd). *Internet of Things (IoT) - Betekenis en definitie*. Opgehaald van Data mining 4U - Weblog over schatgraven in data: <http://www.datamining4u.nl/internet-of-things-betekenis-definitie.php>
- Data.be - Business Info. (2016). *Check 78 in de categorie Bouw van boven- en ondergrondse spoorwegen*. Opgehaald van Data.be: <https://data.be/nl/nace/42120-Bouw-van-bovenen-ondergrondse-spoorwegen>
- Dreamatico.com. (sd). *Railway*. Opgehaald van Dreamatico.com: http://dreamatico.com/data_images/railway/railway-7.jpg
- European Commission. (2010, maart 25). *European Commission*. Opgehaald van List of NACE codes: http://ec.europa.eu/competition/mergers/cases/index/nace_all.html
- Ewer, T. (2015). *Bidsketch*. Opgehaald van 5 steps to preventing scope creep (and still keeping your clients happy): <https://blog.bidsketch.com/clients/preventing-scope-creep/>
- Finler Encyclopedie - Scholto Bos. (2016, november 26). *Innovatie*. Opgehaald van Finler Encyclopedie - financieel-economische begrippen: <https://www.finler.nl/innovatie/>
- Gartner. (2015, augustus 18). *Gartner's 2015 Hype Cycle for Emerging Technologies Identifies the Computing Innovations That Organizations Should Monitor*. Opgehaald van Gartner: <http://www.gartner.com/newsroom/id/3114217>

- Gartner. (2016). *Why Gartner?*. Opgehaald van Gartner: http://www.gartner.com/technology/why_gartner.jsp
- HBO-I . (2015, december 9). *HBO-I Methodekaart*. Opgehaald van HBO-I Hoger beroepsonderwijs en ICT opleidingen.: <http://onderzoek.hbo-i.nl/index.php/Methodenkaart>
- Info Support. (2014, april 26). *Whitepaper Predictive Analytics*. Opgehaald van Info Support: <https://blogs.infosupport.com/media-library/whitepaper-predictive-analytics/>
- Internationaal woordenboek. (sd). *Definitie Machine Learning*. Opgehaald van Internationaal woordenboek: http://internationale-woordenboek.com/definitions/?dutch_word=machine%20learning
- Juice Analytics. (sd). *Chart chooser*. Opgehaald van Juice Analytics Chart Chooser: <http://labs.juiceanalytics.com/chartchooser/index.html>
- Koen van Turnhout, Sabine Craenmehr, Robert Holwerda, Mike Menijn, Jan-Pieter Zwart, René Bakker. (2014, April). *De methodenkaart praktijkonderzoek*.
- Maintenance Assistant. (2016). *Predictive Maintenance (PdM)*. Opgehaald van Maintenance Assistant: <https://www.maintenanceassistant.com/predictive-maintenance/>
- Marketingtermen. (sd). *MoSCoW-methode*. Opgehaald van Marketingtermen.nl: <http://www.marketingtermen.nl/begrip/moscow-methode>
- Microsoft - Azure Data Catalog. (2016). *Microsoft - Azure Data Catalog*. Opgehaald van Microsoft - Azure Data Catalog: <https://azure.microsoft.com/nl-nl/services/data-catalog/>
- Microsoft - Azure Data Factory. (2016). *Microsoft Azure*. Opgehaald van Microsoft - Azure Data Factory: <https://azure.microsoft.com/nl-nl/services/data-factory/>
- Microsoft - Azure Event Hubs. (2016). *Microsoft - Azure Event Hubs*. Opgehaald van Microsoft - Azure Event Hubs: <https://azure.microsoft.com/nl-nl/services/event-hubs/>
- Microsoft - Azure HDInsight (Hadoop). (2016). *Microsoft - Azure HDInsight (Hadoop)*. Opgehaald van Microsoft - Azure HDInsight (Hadoop): <https://azure.microsoft.com/nl-nl/services/hdinsight/>
- Microsoft - Azure Machine Learning. (2016). *Microsoft - Azure Machine Learning*. Opgehaald van Microsoft - Azure Machine Learning: <https://azure.microsoft.com/nl-nl/services/machine-learning/>
- Microsoft - Azure SQL Data Warehouse. (2016). *Microsoft - Azure SQL Data Warehouse*. Opgehaald van Microsoft - Azure SQL Data Warehouse: <https://azure.microsoft.com/nl-nl/services/sql-data-warehouse/>
- Microsoft - Azure Stream Analytics. (2016). *Microsoft - Azure Stream Analytics*. Opgehaald van Microsoft - Azure Stream Analytics: <https://azure.microsoft.com/nl-nl/services/stream-analytics/>
- Microsoft - Cortana Analytics Suite. (2016). *Microsoft - Waarom Cortana Analytics Suite?* Opgehaald van Microsoft - Cortana Analytics Suite: <https://www.microsoft.com/nl-nl/server-cloud/cortana-analytics-suite/why-cortana-analytics.aspx>
- Microsoft. (2016). *Azure IoT Suite*. Opgehaald van Microsoft - Internet of Things: <https://www.microsoft.com/nl-nl/server-cloud/internet-of-things/azure-iot-suite.aspx>
- Microsoft. (2016). *Cortana Analytics Suite*. Opgehaald van Microsoft - Cortana Analytics Suite: <https://www.microsoft.com/en-us/server-cloud/cortana-analytics-suite/what-is-cortana-analytics.aspx>
- Microsoft Azure. (2016). *Azure*. Opgehaald van Microsoft Azure: <https://azure.microsoft.com/en-us/>
- Microsoft- Azure Data Lake. (2016). *Microsoft- Azure Data Lake*. Opgehaald van Microsoft- Azure Data Lake: <https://azure.microsoft.com/nl-nl/solutions/data-lake/>
- Microsoft Cortana. (sd). *Microsoft Cortana*. Opgehaald van Microsoft Cortana: <https://www.microsoft.com/en-us/mobile/experiences/cortana/>
- Microsoft Power BI. (2016). *Microsoft Power BI*. Opgehaald van Microsoft Power BI: www.powerbi.microsoft.com
- Mindtools - Rachel Thompson. (sd). *Stakeholder Analysis - Winning support for your projects*. Opgehaald van Mindtools - Essential skills for an excellent career: https://www.mindtools.com/pages/article/newPPM_07.htm
- MindTools. (2016). *MindTools - Stakeholder matrix*. Opgehaald van https://www.mindtools.com/pages/article/newPPM_07.htm
- Phone Arena - Chris P. (2013, September 16). *Over 21.000 people want Microsoft to keep Cortana, Windows phone's upcoming virtual assistant*. Opgehaald van Phone Arena:

- http://www.phonearena.com/news/Over-21000-people-want-Microsoft-to-keep-Cortana-Windows-Phones-upcoming-virtual-assistant_id47476
- Phone Arena - Chris P. (2013, September 16). *Over 21.000 people want Microsoft to keep Cortana, Windows Phone's upcoming virtual assistant*. Opgehaald van Phone Arena.com: http://www.phonearena.com/news/Over-21000-people-want-Microsoft-to-keep-Cortana-Windows-Phones-upcoming-virtual-assistant_id47476
- PMWIKI.nl. (2016). *Scropecreep*. Opgehaald van Project Management Wiki - Platform Project Management: <http://www.ipma.nl/wiki/kennis/scopecreep>
- Raad van de Europese Economische Gemeenschap. (1990, oktober 9). Invoering NACE-code. *Verordening (EEG) nr. 3037/90 van de Raad van 9 oktober 1990 betreffende de statistische nomenclatuur van de economische activiteiten in de Europese Gemeenschap*.
- Ravenscraft, E. (2015, augustus 3). *Everything you can ask Cortana to do in Windows 10*. Opgehaald van Lifehacker (blog): <http://lifehacker.com/everything-you-can-ask-cortana-to-do-in-windows-10-1721725525>
- Right Markt onderzoek. (sd). *Markt onderzoek - Type onderzoek*. Opgehaald van Right markt onderzoek: <http://www.rightmarktonderzoek.nl/methoden-onderzoek/kwalitatief-onderzoek>
- Strukton Rail - Rolling Stock. (2016). *Strukton On Rail*. Opgehaald van Strukton Rail - Rolling Stock: <http://www.struktonrail.nl/rolling-stock/>
- Swart, N. d. (2010). *Handboek Requirements - De brug tussen business en ICT*. Eburon Business.
- Uitgeverij Couthino . (sd). *Reflectie met de STARR-methode*. Opgehaald van Minor Ondernemen: http://www.minorondernemen.nl/wp-content/uploads/2014/09/extra_info_h3_reflectie_starr-methode.pdf
- Venture Media.nl. (sd). *Gartner Hype Cycle*. Opgehaald van Venture Media.nl - uw gids naar venture capital: <http://www.venturemedia.nl/Gartner-Hype-Cycle>
- WAA Netherlands. (2016). *Wat is Web Analytics*. Opgehaald van Web Analytics: <https://waanetherlands.wordpress.com/wat-is-web-analytics/>
- Welman, B. (sd). *Grafieken, soorten en toepassingen*. Opgehaald van <http://www.softwijs.nl/m-download/m-grafieken/grafieksoorten-pdf?format=raw>
- Werf &. (sd). *Stakeholder matrix: breng de belangen in kaart*. Opgehaald van Werf & over arbeidscommunicatie & recruitment: <https://www.werf-en.nl/nr/nr03/stakeholder-matrix-breng-de-belangen-in-kaart/>
- Werken bij Strukton . (2016). *Werken bij Strukton - Home*. Opgehaald van Werken bij Strukton: <http://www.werkenbijstrukton.nl/n/82/strukton-rail.html>
- Wetenschap.infonu.nl - Rebecca Hoogers. (2011, december 18). *Kwalitatief onderzoek - Experimenteel onderzoek*. Opgehaald van Wetenschap.infonu.nl: <http://wetenschap.infonu.nl/onderzoek/88701-kwantitatief-onderzoek.html>
- Wikipedia. (2016). *Trial and Error*. Opgehaald van Wikipedia.org : <https://nl.wikipedia.org/wiki/Trial-and-error>

Verklarende woordenlijst

Deze lijst omvat alle begrippen met bijbehorende omschrijving die gedurende het document zijn gebruikt.

Begrip:	Omschrijving:
Analytics	Het objectief vastleggen, meten, rapporteren en analyseren van kwantitatieve data met als doel te verbeteren.
Azure (stack)	Windows Azure Platform is een cloud computing-platform van Microsoft waarmee meerdere internetdiensten aangeboden kunnen worden via het internet of binnen de omgeving van het eigen bedrijf. Microsoft wil hiermee de concurrentie aangaan met andere Cloud systemen die software as a service (SaaS) aanbieden, zoals Google App Engine van Google.
Azure Data Catalog	Onderdeel uit de Cortana Intelligence Suite in het Information Management segment van de suite. Azure Data Catalog is een bedrijfsbrede catalogus voor metagegevens en gegevensbrondetectie. Het is een service waarmee gebruikers, analist, wetenschapper en ontwikkelaars op een gebruiksvriendelijke wijze van gegevens kunnen registreren, verrijken, ontdekken, begrijpen en verwerken.
Azure Data Factory	Onderdeel uit de Cortana Intelligence Suite in het Information Management segment van de suite. Deze service dient voor de routing van de data tussen de verschillende onderdelen binnen de Suite. Gebruik de service om gegevens op te nemen uit meerdere bronnen, zowel on-premises als in de Cloud. Met de Azure Data Factory plan en beheer je de gegevenstransformatie en het analyseproces. Zo kunnen met verschillende gegevenspijplijnen taken met de data uitgevoerd worden.
Azure Data Lake	Onderdeel uit de Cortana Intelligence Suite in het Big Data Stores segment van de suite. Azure Data Lake biedt alle mogelijkheden die het voor ontwikkelaars, gegevenswetenschappers en analisten mogelijk maakt om gegevens van verschillende vorm, omvang en snelheid in de Cloud op te slaan, en verwerkingen op uit te voeren op verschillende platformen in verschillende talen. Een data lake kenmerkt zich met een ongestructureerd karakter.
Azure Event Hub	Onderdeel uit de Cortana Intelligence Suite in het Information Management segment van de suite. De Azure Event Hub is een schaalbare publicatie- en aanmeldingsservice die miljoenen gebeurtenissen per seconde kan verwerken en naar meerdere toepassingen kan streamen.
Azure HDInsights (Hadoop)	Een toepassing van het Hadoop Distributed File System (HDFS) binnen Azure met de Cortana Intelligence Suite.
Azure IoT Suite	Met de vooraf geconfigureerde oplossingen die aansluiten op veelvoorkomende Internet of Things-scenario's kunt u tot voorheen onbenutte gegevens vastleggen en analyseren, en zo uw bedrijf drastisch veranderen.
Azure Machine Learning	Onderdeel uit de Cortana Intelligence Suite in het Machine Learning and Analytics segment van de suite. Azure Machine Learning is een element uit

	de Azure Stack waarbij de gebruiker modellen voor predictive analytics kan bouwen, testen en implementeren. De Azure Machine Learning studio biedt een bibliotheek met veel gebruikte algoritmen en voorbeelden waarmee de gebruiker direct aan de slag kan.
Azure SQL Data Warehouse	Onderdeel uit de Cortana Intelligence Suite in het Big Data Stores segment van de suite. Bij traditionele datawarehouses waren altijd vaste combinaties van opslag en rekenprocessen vereist. Hier onderscheidt het door Microsoft aangeboden Azure SQL Datawarehouse zich van de traditionele datawarehouses. Microsoft biedt een datawarehouse waar opslag en rekenprocessen onafhankelijk van elkaar schaalbaar zijn.
Azure Stream Analytics	Onderdeel uit de Cortana Intelligence Suite in het Machine Learning and Analytics segment van de suite. Met Azure Stream Analytics kunnen snelle analyseoplossingen ontwikkeld en geïmplementeerd worden. Hiermee kunnen realtime inzichten verkregen worden van apparaten, sensoren, infrastructuur en toepassingen.
Blob Storage	Online cloud opslag in Azure.
Business Scenario's	Uitkomst uit de Cortana Intelligence Suite. Business Scenario's omvat meerdere onderdelen binnen de suite zoals voorspellingen en aanbevelingen.
Cortana	Persoonlijke spraak gestuurde assistent, ontwikkeld door Microsoft. Het programma is beschikbaar op Windows Phone 8.1, Windows 10 en voor de Microsoft Band. De naam komt van het personage Cortana uit Microsofts Halo-reeks.
Cortana Intelligence Suite	Een samenstelling van diverse volledig geautomatiseerde diensten die geavanceerde analytics mogelijk maken en ieder onderdeel uitmaken van het Microsoft Azure platform. Met als doel data om te zetten in data gestuurde activiteiten.
Datadriven	Beslissingen en acties ondernemen op basis van data(analyse)
Hadoop Distributed File System (HDFS)	Hadoop is een open-source softwareraamwerk van The Apache Software Foundation waarmee computerclusters tot vele Peta bytes aan ongestructureerde data op commodity hardware kunnen verwerken.
Innovatie	Een innovatie kan zowel op technologisch als niet technologisch van aard zijn. Bij technologische innovatie gaat het om vernieuwen of verbeteren van producten, diensten op processen waarmee producten of diensten worden voortgebracht. Bij niet technologische innovaties is er spraken van een vernieuwing in de organisatie
Internet of Things	Afgekort: IoT. Internet of Things (IoT) staat voor het (tijdelijk) verbinden van apparaten met internet om gegevens te kunnen uitwisselen. De verwachting is dat er op den duur meer intelligente apparaten dan mensen op internet zitten. Voorbeelden zijn slimme thermometers, koelkasten, sportkleding met sensoren.
Machine Learning	Een veld van bezighoudt met het ontwerp en de ontwikkeling van algoritmes en technieken die het mogelijk maken computers te leren studeren. Ook wel kunstmatige intelligentie.

Microsoft Visual Studio	Microsoft Visual Studio is een programmeerontwikkelomgeving van Microsoft. Het biedt een complete set ontwikkelingstools om computerprogramma's in diverse programmeertalen voor met name Windows-omgevingen te ontwikkelen. De talen Visual Basic.Net, C#, F# en Visual C++ worden in de standaardeditie meegeleverd.
Multitenancy	Een omgeving waar software of andere diensten draaien op een centrale (vaak cloud) locatie waar onafhankelijke instanties of meerdere applicaties één omgeving delen. De centrale locatie is met logica van elkaar afgeschermd maar fysiek met elkaar geïntegreerd.
Perceptual Intelligence	Uitkomst uit de Cortana Intelligence Suite. Perceptual Intelligence betreft het analyseren van perceptuele gegevens zoals gezichtsuitdrukkingen, zicht, spraak en tekst.
PHP	PHP is een op HTML gebaseerde platformonafhankelijke scriptingtaal die wordt gegenereerd vanuit een webserver. PHP wordt gebruikt om dynamische webpagina's te creëren. Omdat PHP-commando's worden uitgevoerd op de server, kan de gebruiker de PHP-code niet zien.
(Microsoft) Power BI	Microsoft Power BI is een verzameling van verschillende verder doorontwikkelde services uit Microsoft Excel zoals Power Pivot, Power View en Power Query. Deze services bieden functies voor het vinden, analyseren, visualiseren en ontdekkingen kunt delen op intuïtieve wijze met anderen. Daarbij biedt Power BI zowel een desktopapplicatie als een online omgeving waarmee de visualisaties ontwikkeld kunnen worden. Ook zijn de gecreëerde rapporten en dashboard in mobiele apps te benaderen. Binnen de Cortana Intelligence Suite zorgt Power BI voor de visualisatie van de informatie.
Predictive Analytics	Een technologie of toepassing daarvan die het toekomstige gedrag en handelen kan voorspellen op basis van ervaringen en data uit het verleden, met als doel om betere beslissingen te kunnen nemen.
Predictive Maintenance	Het kunnen voorspellen wanneer defecten en/of problemen optreden om hier tijdig op in te kunnen spelen en op deze wijze de inzetbaarheid te waarborgen. Het planmatig data gedreven in kunnen plannen van onderhoud.
Scope creep	Scope creep is het fenomeen waarbij de scope langzaam en bijna ongemerkt wordt uitgebreid zonder dat een nieuwe scope expliciet is vastgesteld.
SQL (statement)	Structured Query Language. Een gestructureerde taal waarin commando's aan databases gegeven kunnen worden met een standaard gedefinieerde opbouw.
Tenant	Een tenant is een groep gebruikers met dezelfde toegangsrechten tot bepaalde software.
Triangulatie	Het combineren van verschillende theorieën, methoden en bronnen om tot een beter en vollediger antwoord op een vraag te komen.

Tabel 5, Verklarende woordenlijst

Lijst van figuren en tabellen

Lijst van figuren

Figuur 1, Logo Centric	11
Figuur 2, Kernactiviteiten Centric	12
Figuur 3, Logo Strukton Rail	14
Figuur 4, Gartner Hype cycle	15
Figuur 5, Word Cloud Innovation	18
Figuur 6, Spoorweg	23
Figuur 7, Voordelen Cortana Intelligence Suite	26
Figuur 8, Elementen in de Cortana Intelligence Suite	27
Figuur 9, Gartner's magic quadrant for business intelligence and analytics	28
Figuur 10, Stakeholdermatrix Strukton Rolling Stock	31
Figuur 11, Algemene stakeholdermatrix	33
Figuur 12, Schematische weergave Cortana Intelligence Suite	38
Figuur 13, Voorbeeld .XML-data Strukton Rolling Stock	39
Figuur 14, Datamodel Azure SQL-database	42
Figuur 15, Cortana Intelligence Suite - Solution Template	47
Figuur 16, SQL-query aanmaken tabellen en relaties in Visual Studio	49
Figuur 17, Demo dashboard Strukton Rolling Stock	53
Figuur 18, Operationeel dashboard registratie data Strukton Rolling Stock	54
Figuur 19, Solution templates uit de Cortana Intelligence Gallery	59

Lijst van tabellen

Tabel 1, Versiegeschiedenis	5
Tabel 2, Betekenis variabelen Strukton remote monitoring data	41
Tabel 3, MoSCoW indeling informatiebehoeften	45
Tabel 4, Review op MoSCoW indeling informatiebehoeften	58
Tabel 5, Verklarende woordenlijst	70
Tabel 6, Bijlagen	72

Bijlagen

Bijlage nummer:	Naam bijlage:	Datum bijlage:	Versie Bijlage:
I	Afstudeerplan	2 februari 2016	1.0
II	Gantt Chart afstudeerplan	12 februari 2016	1.0
III	Business Model Canvas – Strukton Rolling Stock	18 februari 2016	1.0
IV	Toelichting op de Gartner hype cycle	-	-
V	Toelichting op de HBO-I onderzoeksmethode	-	-
VI	NACE-codes	-	-
VII	Onderzoeksplan	16 februari 2016	1.2
VIII	Gespreksverslag Boudewijn van der Zwan	22 februari 2016	1.0
IX	Gespreksverslag Peter van Klaveren	16 februari 2016	1.0
X	Omschrijving Cortana Intelligence Suite	-	-
XI	Gespreksverslag Han Geerligs	17 februari 2016	1.0
XII	Gespreksverslag Marcel van der Linden	17 februari 2016	1.0
XIII	Toelichting op de stakeholdermatrix	-	-
XIV	PHP-script data import in SQL-database	-	-
XV	Schermafbeeldingen Power BI-dashboard demo omgeving	24 mei 2016	1.0
XVI	Schermafbeeldingen Power BI-dashboard operationele omgeving	24 mei 2016	1.0
XVII	Visiedocument Centric t.a.v. SaaS	-	-
XVIII	Voortgangsrapportage	6 april 2016	1.0
XIX	Competentieverantwoording	-	-
XX	Evaluatieformulier Afstuderen tbv Haagse Hogeschool	30 mei 2016	1.0

Tabel 6, Bijlagen