



Rijkswaterstaat
Ministerie van Infrastructuur en Waterstaat |

DATAMODELLEREN VOOR IV-PROJECTEN

AFSTUDEERVERSLAG

Rijkswaterstaat

Auteur: Bo-Lan Schouten
Stageperiode: 01 februari 2022 – 30 juni 2022
Stagebedrijf: Rijkswaterstaat
Opleiding: Informatica, Hogeschool Leiden
Versienummer: 2.0
Versietype: Definitieve versie
Versiedatum-aanduiding: 14-06-2022

VOORWOORD

Mijn afstudeerstage heb ik gedaan bij Rijkswaterstaat. De afgelopen 5 maanden waren voor mij zeer waardevolle en leerzame maanden. Hierbij heb ik mij voornamelijk beziggehouden met het modelleren van data en advies geven op het onderhouden van de data. Hierbij heb ik dus ook veel kennis opgedaan van verschillende data management frameworks en UML. Daarnaast heb ik ook de mogelijkheid gekregen om te werken met Sparx Enterprise Architect. Ik heb hier veel van geleerd, zowel over de tooling als over het modeleren binnen deze omgeving. Daarnaast heb ik ook veel kennis op kunnen doen van data management frameworks en deze uitgebreid kunnen analyseren. Al met al kan ik zeggen dat ik dit een zeer leerzame stage vond, waarbij ik nieuwe kennis en ervaring heb opgedaan.

Graag bedank ik mijn stagebegeleiders, Rob Smits en Rudolf Brugman. Gedurende mijn stage heb ik veel aan hen gehad. Bij problemen kon ik altijd bij hen terecht. Ook wil ik graag alle stakeholders bedanken voor hun tijd en de informatie die zij aan mij konden verstrekken. Extra documenten die mij eventueel konden helpen waren zeer waardevol voor het resultaat en zonder hen was ik waarschijnlijk niet zelf op alle documenten gekomen.

SAMENVATTING

Rijkswaterstaat is verantwoordelijk voor de hoofdinfrastructuur en hoofdwatersystemen binnen Nederland. Het aanleggen en onderhoud van (vaar)wegen, managen van water. Het bouwen van bijvoorbeeld een brug wordt binnen Rijkswaterstaat gezien als een project en van zo'n project worden (veel) data vastgelegd. Denk hierbij bijvoorbeeld aan wie hiervoor verantwoordelijk is, in wat voor projectfase het zich verkeerd, de startdatum, etc. Momenteel zijn deze modellen vastgelegd in Microsoft Word wat er onder andere voor zorgt dat de structuur onoverzichtelijk is en moeilijk is bij te werken..

Het doel van de stage is om UML modellen op te stellen voor IV-projecten en deze in Sparx Enterprise Architect (hierna te noemen 'EA') vast te leggen. EA is de architectuurtool die binnen Rijkswaterstaat gebruikt wordt. Door dit in schematisch in EA vast te leggen worden definities duidelijk en wordt het inzichtelijk welke klasse en relatie heeft met welke klasse verbonden is. Hierdoor worden dashboard ook beter. Daarnaast moet er ook advies gegeven worden over hoe deze data(modellen) het beste onderhouden/beheerd kunnen worden.

Dit advies beveelt een management framework aan die het best bij Rijkswaterstaat past. Om uiteindelijk tot datamodellen en een advies te komen is er door mij onderzoek gedaan.

Dit document beantwoordt dan ook de vraag: Hoe kan er vanuit de wensen van Rijkswaterstaat nieuwe verbeterde datamodellen gemaakt worden ten aanzien van IT-Projecten volgens onderbouwde literatuur en hoe kunnen deze datamodellen het best beheerd worden? En gaat in op de Rijkswaterstaat zelf, UML, kennisgebieden binnen DAMA DMBOK en er wordt gekeken naar de functionaliteiten binnen EA. Door gebruik te maken van verschillende onderzoeksmethodes, denk hierbij aan literatuuronderzoek, interviews, documentanalyses is dit onderzocht en uiteindelijk is er een antwoord op de vraag gekomen.

INHOUDSOPGAVE

VOORWOORD.....	2
SAMENVATTING	3
FIGURENLIJST	6
TABELLENLIJST.....	6
BEGRIPPENLIJST.....	7
1. INLEIDING.....	8
2. BESCHRIJVING STAGEORGANISATIE.....	9
2.1 Stagebedrijf.....	9
2.1.1 Organogram.....	9
2.2 Stagebegeleiding.....	10
2.3 Stakeholders	10
3. PROBLEEMSTELLING.....	12
4. DE CONTEXT	13
4.1 Het proces.....	13
4.2 De data architectuur	15
4.2.1 FAIR principes.....	16
5. HET ONDERZOEK	17
6. METHODOLOGIE	18
7. ONDERZOEKSRESULTATEN – DEELVRAGEN.....	20
7.1 Is EA geschikt voor het modelleren van de gewenste UML-modellen?.....	20
7.2 Wat zijn de requirements waar de modellen aan moeten voldoen?	21
7.3 In hoeverre is door het bedrijf dit probleem omtrent modelleren van conceptuele datamodellen al behandeld?.....	23
7.4 Met welke ‘regels’ moet er rekening worden gehouden bij het modelleren van een UML Klassendiagram?	26
7.5 Wat zegt DAMA DMBOK over het onderhouden/beheren van data(modellen) ?	29
7.6 Wat zegt De Cloud Data Management Capabilities Framework over het onderhouden/beheren van data(modellen) (Hier naar kijken)!	34
8. BEROEPSPRODUCT	36
8.1 De datamodellen.....	36
8.2 Het advies	39

8.2.1	Datamodellen ontwerp.....	39
8.2.2	Data management framework.....	41
9.	REFLECTIE.....	44
9.1	Competenties.....	44
9.1.1	A-Competenties.....	44
9.1.2	B-Competenties.....	46
9.2	Planning.....	48
9.3	Uitkomsten.....	49
9.4	Begeleiding.....	49
10.	BIJLAGE.....	51
10.1	Bijlage A: Literatuurlijst.....	51
10.1.1	Afbeeldingen.....	51
10.1.2	Websites.....	51
10.1.3	Boeken.....	52
10.1.4	PDF-documenten.....	53
10.2	Bijlage B: Het bedrijfsobjectmodel RWS (BOMR).....	54
10.2.1	Het huidige bedrijfsobjectmodel RWS (BOMR).....	54
10.2.2	Het bedrijfsobjectmodel RWS (BOMR) 2009.....	54
10.3	Bijlage C: Interviews.....	55
10.3.1	Interview/vragenlijst Victor van der Kloet.....	55
10.3.2	Interview Peter Bernhard.....	56
10.3.3	Interview Rob Smits.....	66
10.4	Bijlage D: Workshop Sparx en UML.....	73
10.5	Bijlage E: Functionaliteiten per editie Sparx Enterprise Architect.....	74
10.6	Bijlage F: Afstudeervoorstel.....	81
10.7	Bijlage G: Adviesrapport Dataframework en datamodellen.....	96

Versie geschiedenis

Tabel 1: Versiebeheer

Versie	Datum	Status	Opmerkingen/gewerkt aan
0.1	12-04-2022	Ontwikkeling	Opmaak, Stage organisatie, Methodologie, Probleem, Onderzoek
0.2	18-04-2022	Ontwikkeling	Begrippenlijst, competenties
0.3	21-04-2022	Ontwikkeling	BPMN, context aan de opdracht, aanpak van de onderzoeksresultaten, beroepsproduct
0.4	02-05-2022	Pre-final	Feedback van intervisie verwerkt
1.0	14-05-2022	Pre-final	Onderzoeksresultaten, context, beroepsproduct, reflectie
1.1	02-06-2022	Pre-final	Laatste dingen bijwerken.
2.0	03-06-2022	Definitief	

FIGURENLIJST

FIGUUR 1: ORGANOGRAF RIKSWATERSTAAT (RIKSWATERSTAAT, 2021c)	10
FIGUUR 2: ORGANOGRAF CIV (RIKSWATERSTAAT, 2022)	10
FIGUUR 3: BPMN MODEL	13
FIGUUR 4: USE CASE DIAGRAM	14
FIGUUR 5: SAMENHANG DATA-ARCHITECTUUR (RIKSWATERSTAAT, 2021b)	15
FIGUUR 6: SCHEMATISCHE WEERGAVE AANPAK STAGE	19
FIGUUR 7: VOORBEELD ASSOCIATION NAMEN (BOOCH & RUMBAUGH, 2005a)	27
FIGUUR 8: VOORBEELD ASSOCIATION ROLLEN (BOOCH & RUMBAUGH, 2005b)	27
FIGUUR 9: VOORBEELD DEPENDENCY RELATIE (BOOCH & RUMBAUGH, 2005c)	28
FIGUUR 10: VOORBEELD GENERALISATION RELATIE (BOOCH & RUMBAUGH, 2005d)	28
FIGUUR 11: VOORBEELD AGGREGATION RELATIE	28
FIGUUR 12: VOORBEELD COMPOSITION RELATIE	29
FIGUUR 13: VOORBEELD MULTIPLICITY (G ENOVA ET AL., 2002, FIGUUR 1)	29
FIGUUR 14: HET DATAMODEL BEDRIJFSVOERING IV	36
FIGUUR 15: HET DATAMODEL IV-PROJECTEN	37
FIGUUR 16: HET DATAMODEL IV-LEVERPROJECT	38
FIGUUR 17: HET DATAMODEL IV IN GWW-PROJECT	38
FIGUUR 18: HET DATAMODEL IV-ONTWIKKELPROJECT	39
FIGUUR 19: OBJECTMODEL RWS (KAMP, 2009)	55
FIGUUR 20: BEWIJS VAN DEELNAME WORKSHOP	73

TABELLENLIJST

TABEL 1: VERSIEBEHEER	5
TABEL 2: STAKEHOLDERS ANALYSE	11
TABEL 3: OVERZICHT ONDERZOEKSMETHODES	18
TABEL 4: MoSCoW ANALYSE	22
TABEL 5: GAP-ANALYSE UITSTRALING	24
TABEL 6: GAP-ANALYSE BETROUWBAARHEID	24
TABEL 7: GAP-ANALYSE ZEKERHEID	25
TABEL 8: MULTIPLICITY	29
TABEL 9: ENTERPRISE ARCHITECTUUR DOMEINEN (DMBOK2, 2017)	33
TABEL 10: L-MATRIX REQUIREMENTS DAMA DMBOK vs CDMC	43
TABEL 11: AANTONEN B-COMPETENTIE OPDRACHT IN CONTEXT	46
TABEL 12: AANTONEN B-COMPETENTIE ADVISEREN DATAOPLOSSING	47
TABEL 13: AANTONEN B-COMPETENTIE ANALYSEREN DATA ARCHITECTUUR	48
TABEL 14: ENTERPRISE ARCHITECTUUR DOMEINEN (DMBOK2, 2017)	105
TABEL 15: L-MATRIX REQUIREMENTS DAMA DMBOK vs CDMC	108

BEGRIPPENLIJST

<i>Begrip</i>	<i>Betekenis</i>
Attribuut	Een eigenschap van een klasse, bijvoorbeeld een projectnummer of een status
BOMR	Bedrijfsobjectmodel Rijkswaterstaat
CIV	Centrale informatie voorziening, dienst binnen Rijkswaterstaat
CMD	Een conceptueel datamodel
CMDB	Configuratie management database, Database met alle componenten van het IV-landschap
Directie	Een directie bevat meerdere afdelingen
DT	Directie team (medewerkers)
EA	Enterprise Architect, hiermee wordt Sparx Enterprise Architect bedoeld.
IGA	Inwinning & Gegevens analyse (directie van de CIV)
Informatie voorziening	Proces om de verzamelde informatie beschikbaar te stellen
IenW	Infrastructuur en waterstaat (ministerie)
IRN	IV Rijkswaterstaat netwerken (directie van de CIV)
IV	Informatievoorziening
IV-projecten	Informatievoorzieningsprojecten, dit is een geheel van samenhangende activiteiten gericht op ontwikkeling of vernieuwing, die gecoördineerd in een tijdelijke organisatie worden uitgevoerd, teneinde geformuleerde doelstellingen te realiseren binnen vastgestelde kaders: scope, tijd, kwaliteit en budget. Een project levert één of meerdere producten op, die door gebruikers in gebruik worden genomen.
MT	Management Team (afdelingshoofden)
OSR	Ontwikkeling & Services (directie van de CIV)
OTL	Object Type Library, Bedrijfsbreed informatie model/Rijkswaterstaat brede conceptuele informatie model van de objecten; bruggen, sluizen, stuwen.
PPO	Programma's, projecten en onderhoud (dienstonderdeel RWS)
PID	Project initiatie document (projectplan)
Relatie	Een associatie tussen verschillende klassen. Bijvoorbeeld een project heeft een projectmanager (relatie), dan is er een relatie tussen project en medewerker
RWS	Rijkswaterstaat (organisatie)
SAS	Statistical Analysis System, software ontwerpen door SAS institute voor het analyseren van data en het rapporteren ervan.
S&B	Strategie & beleid (afdeling bij de CIV)
TP	Technische projecten (afdeling bij de CIV)

1. INLEIDING

Data wordt steeds belangrijker en er komt ook steeds meer data bij. De hoeveelheid data blijft ook groeien, waar het in 2010 nog ging om 2 zettabytes (ZB) aan data, was dit in 2020 64 ZB aan data. In 2021 was er wereldwijd 79 ZB aan data. De hoeveelheid data in de wereld blijft ook exponentieel hard groeien. De verwachting voor 2025 is dat er dan wereldwijd 181 ZB aan data bestaat. Dat wil zeggen dat er in 2025 181 biljoen gigabyte aan data ter wereld is (Statista Research Department, 2022).

Met data kunnen verwachtingen voorspeld worden, denk hierbij aan het weer, wanneer objecten onderhouden of vervangen moeten worden en hoeveel geld hiervoor nodig is. Ook voor bedrijven zijn data handig. Dit kan gaan om data van hen klanten/gebruikers, Albert Heijn die de koop gegevens van klanten onthoudt of Google die een heel gebruikersprofiel opstelt waarbij het je leeftijd, geslacht, opleidingsniveau, et cetera. inschat. Maar dit kan ook data betreffen die vrijkomt bij bedrijfsprocessen: duur van een project, benodigd budget, resultaten, et cetera. Data is belangrijk om te kunnen sturen op de bedrijfsprocessen en te kunnen verantwoorden wat er met de inzet van geld en mensen is gerealiseerd.

Ook Rijkswaterstaat slaat data op. Data over wegen, water, projecten, et cetera. Alleen hierbij treden problemen op. Soms is het onduidelijk wat voor soort data opgeslagen moet worden of welke data opgeslagen zijn, hoe dit allemaal met elkaar verbonden is en wat er precies mee wordt bedoeld. Het probleem dat hieruit volgde was weer dat dashboard niet optimaal waren. De oplossing voor dit probleem: Datamodellen. Mijn afstudeeropdracht was dan ook om data van IV-projecten te modelleren volgens UML in EA en vervolgens advies te geven over het onderhouden van data(modellen).

Dit document gaat in op het verloop van de stage. In eerste instantie wordt er gekeken naar Rijkswaterstaat zelf, wie is het en wat doet het. Vervolgens wordt het onderzoek geïntroduceerd en wordt het onderzoek beschreven. Zo wordt er gekeken of EA ook echt de juiste tooling is, komen de requirements waar de datamodellen aan moeten voldoen naar voren, in hoeverre het bedrijf al gewerkt heeft aan het probleem, welke 'regels' er binnen UML zijn en twee kennisgebieden binnen data management frameworks worden uitgelegd (DAMA DMBOK/Clouw Data Management Capabilities, waarom er hiervoor gekozen is wordt ook uitgelegd). Uiteindelijk komen er antwoorden tot stand wat moet leiden tot (correcte) datamodellen en een advies. Tot slot zal dit hele proces geëvolueerd worden.

Het doel van dit document is dan ook om door het verloop van de stage uit te leggen, en zo aan te tonen dat ik beschik over de A-competenties en de 3 gekozen B-competenties. De gekozen B-competenties zijn als volgt:

- Organisatieprocessen analyseren
- Organisatieprocessen adviseren
- Organisatieprocessen adviseren.

2. BESCHRIJVING STAGEORGANISATIE

2.1 Stagebedrijf

Rijkswaterstaat houdt zich bezig met Infrastructuur en waterstaat. Het is verantwoordelijk voor beheren van grote wateren maar ook voor het onderhouden van stormvloedkeringen, stuwen en dijken. Daarnaast is Rijkswaterstaat ook verantwoordelijk voor het onderhoud en de kwaliteit van de rijkswegen. Het is dan ook de uitvoeringsorganisatie van het ministerie van Infrastructuur en Waterstaat (Ministerie van Infrastructuur en Waterstaat, 2021b).

De missie luidt dan ook als volgt: *“Rijkswaterstaat is de uitvoeringsorganisatie van het ministerie van Infrastructuur en Waterstaat. We beheren en ontwikkelen de rijkswegen, -vaarwegen en –wateren en zetten in op een duurzame leefomgeving. Samen met anderen werken we aan een land dat beschermd is tegen overstromingen. Waar voldoende groen is, en voldoende en schoon water. En waar je vlot en veilig van A naar B kunt. Samen werken aan een veilig, leefbaar en bereikbaar Nederland. Dat is Rijkswaterstaat.”* (Ministerie van Infrastructuur en Waterstaat, 2021c)

De organisatie heeft verschillende visies; een marktvisie, netwerkbeheervisie maar ook een visie op informatievoorziening. Waarbij de visie op informatievoorziening als volgt is: *“Om goede samenwerking mogelijk te maken en onze Rijkswaterstaatopgave succesvol aan te kunnen gaan, is het ontsluiten, onderhouden, beheren en verspreiden van informatie, maar ook de toegankelijkheid, bruikbaarheid en betrouwbaarheid van deze informatie van cruciaal belang. Net als de applicaties en de infrastructuur die nodig zijn om die informatie in te winnen, te beheren en te ontsluiten. Een efficiënte en effectieve informatiehuishouding die past bij de opgave en werkwijze van Rijkswaterstaat is dan ook cruciaal. Informatievoorziening vormt de ruggengraat van onze dagelijkse operatie”* (Rijkswaterstaat, 2021d, p. 4)

Rijkswaterstaat kent zeven landelijke en zeven regionale organisatieonderdelen. Mijn stage vond plaats bij de dienst Centrale Informatie voorziening (CIV). Zij houden zich voornamelijk bezig met de ontwikkeling en beschikbaarheid van alle informatie. Zo zorgen ze enerzijds voor industriële automatisering bij bijvoorbeeld bruggen en tunnels, anderzijds houden ze zich ook bezig met kantoorautomatisering. Ze leveren programma's, producten en diensten voor alle onderdelen binnen Rijkswaterstaat (Ministerie van Infrastructuur en Waterstaat, 2021a).

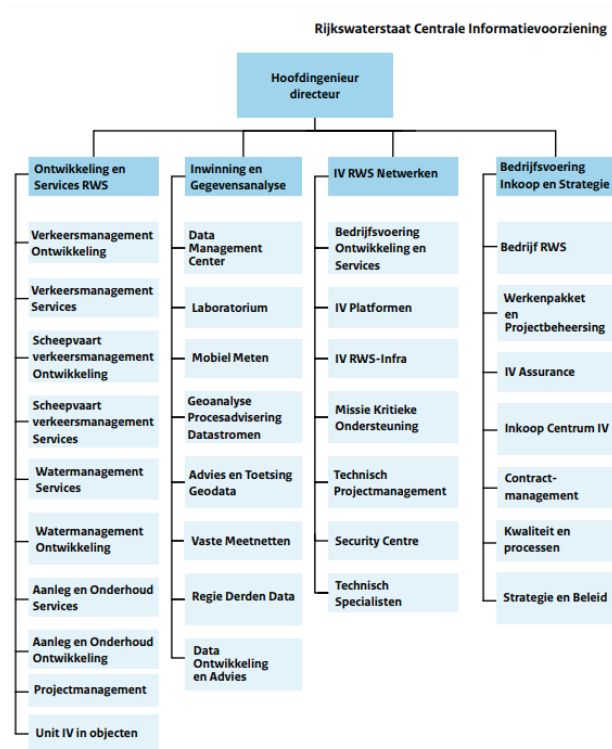
De Centrale informatievoorziening bestaat uit verschillende directies. Hierbij val ik onder de directie: Bedrijfsvoering Inkoop en Strategie en binnen de directie behoor ik tot de afdeling Strategie en Beleid. Deze afdeling ondersteunt en adviseert het directieteam van de CIV, de hoofdingenieur-directeur en de Chief Information Officer (CIO) ofwel proceseigenaar informatievoorziening in hun taak: *“de besturing van en richting geven aan de Informatievoorziening voor Rijkswaterstaat”*. (Rijkswaterstaat, 2020). De datamodellen dragen dan ook bij aan het bouwen van dashboards. Deze dashboards helpen bij het ondersteunen en adviseren van het directieteam.

2.1.1 Organogram

Het organogram van Rijkswaterstaat ziet er als volgt uit:



Figuur 2: Organogram Rijkswaterstaat (Rijkswaterstaat, 2021c)



Figuur 1: Organogram CIV (Rijkswaterstaat, 2022)

2.2 Stagebegeleiding

Gedurende mijn stage ben ik twee dagen per week op kantoor Rijkswaterstaat te Delft geweest. Hierdoor werd het makkelijker om met elkaar te communiceren omdat je bij elkaar in de buurt zit. Daarnaast was er op de donderdag een dag start, dit was met de hele afdeling en hierbij vertelt iedereen wat hij/zij die dag gaat doen. Ook had ik een wekelijkse meeting met Victor van der Kloet, met wie ik nauw samenwerkte. Victor van der Kloeten was een informatieanalist, een van zijn verantwoordelijkheden waren de datamodellen met betrekking tot IV-projecten. Tijdens mijn gesprekken met hem werd er vooral gekeken naar de voortgang, werd er gekeken naar de datamodellen en of ik ergens problemen ondervond.

Helaas veranderde Victor halverwege mijn stage van baan waardoor dit uiteindelijk niet meer kon. Omdat ik in eerste instantie vooral voortgangsgesprekken met hem had werd dit overgenomen door mijn bedrijfsbegeleider. Ik had wekelijks een gesprek met Rob Smits waarbij er gekeken werd naar de algemene voortgang en of de stage nog steeds goed ging. Mijn bedrijfsbegeleider kon ik altijd via teams of per mail te bereiken.

2.3 Stakeholders

Bij mijn stage heb ik te maken met volgende stakeholders. De informatiemanager van het proces informatievoorziening (Rob Smits) en de informatieanalist met wie ik nauw samenwerk (Victor van der Kloet). Daarnaast hebben ook data architecten en adviseurs mij geholpen bij het onderzoek. Deze mensen hebben mij voornamelijk geholpen bij het begrijpen van de organisatie zelf. Hoe bijvoorbeeld projecten in elkaar zitten en wat zij nodig hebben aan data. Maar ook om gewoon de gegevens te begrijpen, bijvoorbeeld bepaalde termen, boden zij hulp waar nodig of bij definitieproblemen hebben zij opheldering gegeven.

Als ik kijk naar de stakeholders die het product gaan gebruiken dan zijn dit de medewerker die dashboards bouwen. Met deze mensen heb ik echter niet zelf mee te maken gehad. Een overzicht van de stakeholders waar ik mee te maken had is hieronder te vinden.

Tabel 2: Stakeholders analyse

<i>Naam</i>	<i>Functie</i>	<i>Intern/Extern</i>	<i>Primair/Secundair</i>	<i>Belang & impact</i>
<i>Rob Smits</i>	Informatie Manager	Intern	Primair	5
<i>Victor van der Kloet</i>	Informatieanalyst	Intern	Primair	5
<i>Mieke van Nierop</i>	Informatieanalyst	Intern	Primair	4
<i>Margreet Kampen</i>	Data Architect	Extern	Secundair	4
<i>Renee Mantel</i>	Applicatiemanager/functioneel beheerder IV-tooling	Extern	Secundair	3
<i>Arthur de Waard</i>	Clustercoördinator Cluster IV-leverprojecten	Extern	Secundair	3
<i>Peter Bernhard</i>	Senior adviseur - Domein architectuur	Extern	Secundair	3
<i>Edwin van 't Walderveen</i>	Data architect	Extern	Secundair	2
<i>Henk van Evert</i>	Senior-adviseur Kwaliteit en Processen	Intern	Primair	2
<i>Rutger van Beek</i>	Adviseur Risicomanagement	Intern	Primair	2
<i>Jo Nijssen</i>	Externe Medewerker	Extern	Secundair	2
<i>Wim Ter Borg</i>	Adviseur productieprocessen IGA	Extern	Secundair	2

Hierbij is gekeken naar in welke mate ze betrokken zijn bij de opdracht, of zij een groot belang hierbij hebben en hoe groot hen impact was.

- **Interne stakeholders** zijn nauw betrokken bij de opdracht en hebben ook bijgedragen/dragen bij aan de opdracht.
- **Externe stakeholders** hebben niet meegewerkt aan de opdracht en staan er ook iets verder van af.
- **Primaire stakeholders** hadden een (grote) belang op de opdracht.
- **Secundaire stakeholders** hebben een iets minder groot belang bij de opdracht.
- Op basis van een score is er aangegeven hoe groot het belang/impact de opdracht had op de stakeholders. Hierbij geldt: 1 is helemaal niet, en 5 helemaal wel.

Om de score te bepalen is er gekeken naar de hoeveelheid impact/betrokkenheid van de personen bij mijn stage. 5 houdt in zij waren nauw betrokken met de opdracht en hadden dus ook veel impact op de uitkomst en belang bij het resultaat. Een score van 4 houdt in dat ze niet nauw betrokken waren met de opdracht maar hen input wel erg belangrijk was. Ook heeft het resultaat niet directe invloed op hen. Een score van 3 geeft aan dat ze niet nauw betrokken waren bij de opdracht, maar hen input wel relevant was voor de opdracht zelf, maar in mindere mate dan stakeholders met een score van 4. Een score van 2 geeft aan dat ze vrijwel niet of nauwelijks betrokken waren bij de opdracht en ook niet veel impact hadden op de opdracht. Zij droegen voornamelijk bij aan het beter begrijpen van de context.

3. PROBLEEMSTELLING

Rijkswaterstaat heeft verschillende conceptuele datamodellen, deze staan echter in Microsoft Word bestanden waardoor dit momenteel onoverzichtelijk en moeilijk te onderhouden is. Deze 'modellen' zijn namelijk uitgelegd via tabellen. Er is aangegeven wat de klassen zijn, welke attributen en relaties deze hebben en vervolgens staat per attribuut de definitie. Hier is geen schematisch weergave van en het is ook niet in één oogopslag te zien welke klassen een relatie hebben met andere klassen. Daarnaast is het ook moeilijk om relaties tussen verschillende modellen weer te geven. Mogelijk kan er een verwijzing in staan naar een ander model, maar men moet dan weer in alle Microsoft Word bestanden gaan zoeken naar het desbetreffende model. Dit is vooral voor de mensen die de dashboards bouwen zeer lastig. Door dit schematische weer te geven in datamodellen krijgen zij inzicht in de samenhang en eenduidigheid in de gegevens. De conceptuele datamodellen dienen als eerste stap na het uitvragen van informatiebehoefte om uiteindelijk tot een dashboard te kunnen komen.

Er mist binnen de data ook uniformiteit. Met bepaalde zaken bedoelt de ene persoon A en voor de ander persoon betekent dit B. Dit zorgt er ook voor dat de data ook niet van optimale kwaliteit is. Dashboards kunnen zo bijvoorbeeld de verkeerde data laten zien.

De oplossing voor bovenstaande problemen is om de datamodellen in EA te implementeren. Door het onder andere in UML vast te leggen kunnen de definities vast worden gelegd. Er kunnen duidelijke eisen gedefinieerd worden waaraan een klasse aan moet voldoen en er kan ook gezien worden waar zo'n klasse uit bestaat. Door gebruik te maken van EA kunnen de datamodellen in kaart worden gebracht en zorgt dit ervoor dat het beter te onderhouden/bij te werken is. Alles staat dan namelijk in dezelfde omgeving.

Ik zal kijken naar de huidige stand van zaken. Hoe zien de huidige conceptuele datamodellen eruit en hoe gaat Rijkswaterstaat er nu mee om. Vervolgens zal ik het omzetten naar een UML-model en in EA zetten. Hierbinnen kunnen dan de eerder opgestelde definities van bepaalde objecten vastgelegd worden. Binnen Rijkswaterstaat is er gekozen om EA te gebruiken, hier hebben ze ook een licentie op en moet dus gebruikt worden. Rijkswaterstaat heeft voor deze tooling gekozen omdat het een veel gebruikte tool is door andere grote bedrijven en goedkoper is dan Troux. Troux was de modelleertool die hiervoor gebruikt werd door Rijkswaterstaat. Dit wordt verder uitgewerkt in [hoofdstuk 7.1](#). Dit is dus vanuit het stagebedrijf besloten en hier heb ikzelf geen invloed op.

4. DE CONTEXT

4.1 Het proces

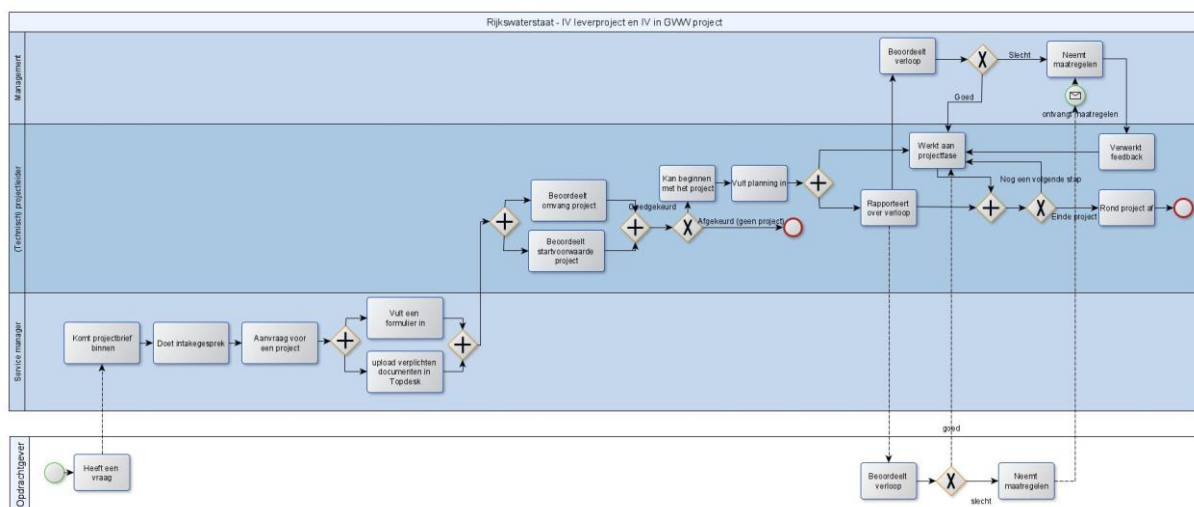
Dat er onduidelijkheid is over wat er wordt bedoeld met begrip A waardoor bijvoorbeeld dashboards verkeerde resultaten tonen komt nu naar boven omdat er op het moment dashboards gebouwd worden die de data direct uit databases halen. Voorheen werden overzichten handmatig opgebouwd uit spreadsheets, waarbij gegevens ook direct werden gecorrigeerd. Het probleem komt dus ook naar boven bij het bouwen van dashboards voor het management.

Zoals in de interviews naar voren is gekomen, komt dit door de 'communicatie' tussen verschillende afdelingen. Binnen één afdeling is het niet zo'n probleem, maar wanneer verschillende afdelingen met elkaar gaan samenwerken en zij binnen hun domein verschillende definities hanteren voor hetzelfde begrip ontstaat er een soort van miscommunicatie. Het ene domein bedoelt dit met dat begrip en het andere domein bedoelt weer iets heel anders. Rijkswaterstaat wil meer een data gedreven organisatie worden waardoor er dus ook betere rapportage aan het management, directieteams en Bestuur RWS geleverd kunnen worden. Daarom dat enerzijds data steeds belangrijker wordt en anderzijds, dat afdelingen sneller geneigd zijn op samen te werken en hen data te delen. Hierbij trad dus het probleem op dat niet elk domein met begrip A hetzelfde bedoeld. Daarom dat Rijkswaterstaat nu bezig is met het opstellen en visualiseren van de datamodellen, dit helpt dus ook bij het opstellen van deze dashboards.

De datamodellen gaan over IV-project. Rijkswaterstaat kent drie verschillende type projecten en die worden als volgt gedefinieerd:

- IV-leverproject: het projectmatig leveren van beheerde standaard IV-producten of diensten.
- IV-ontwikkelproject: Ontwikkelen van nieuwe of vernieuwde producten en diensten binnen CIV.
- IV in GWW (Grond- Weg en Waterbouw) project: Leveren van IV-diensten in grote infrastructurele projecten

Als er gekeken wordt naar het proces/verloop van een project ziet het versimpelde model er als volgt uit. Hierbij is er voornamelijk gekeken naar waar de data vrijkomt (de momenten dat de data wordt vastgelegd):



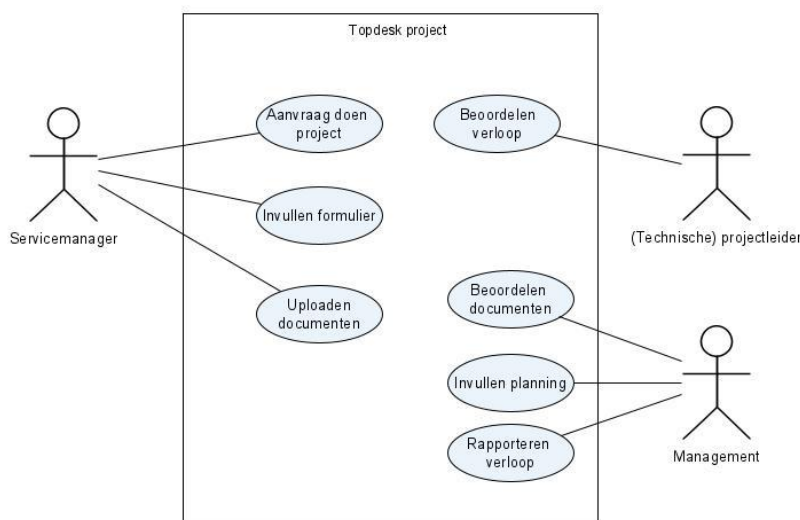
Figuur 3: BPMN model

*In bovenstaande BPMN model heeft alleen specifiek betrekking op IV-Leverprojecten en IV in GWW projecten. Het proces van een ontwikkelproject wijkt af van de ander twee projecten maar als er puur gekeken wordt naar de momenten waarbij data vrijkomt, is het bovenstaand proces nog steeds van toepassing. De manier hoe de data wordt vastgelegd is in dat opzicht hetzelfde. Er wordt nog steeds in topdesk bijgehouden en gerapporteerd over hoe een project op basis van kwaliteit, geld en tijd verloopt.

Als eerste begint het bij de opdrachtgever. Hij heeft een vraag en gaat hiermee naar de servicemanager. Op basis hiervan heeft de servicemanager een intakegesprek met de opdrachtgever. Vervolgens doet de servicemanager een aanvraag voor een project. Dit doet hij via de Topdesk omgeving. Hij moet een formulier invullen waarbij het project gedefinieerd wordt. Hierbij wordt gekeken is het ook echt een project is en geen klus. Dit wordt bepaald aan de hand van een score. Daarnaast moet er dus ook een opdrachtgever zijn, moeten algemene zaken van een project gedefinieerd worden (denk hierbij aan begindatum, einddatum, workload, opdrachtgever, etc.) en er moeten documenten opgeleverd worden. Zodra dit allemaal gedaan is wordt de aanvraag beoordeeld. Dit wordt gedaan op basis van de startvoorwaarden en de omvang van het project. Wanneer dit wordt afgekeurd wordt het proces beëindigd. Wordt dit alles goedgekeurd gaat het proces verder. Er kan namelijk begonnen worden aan het project.

De projectleider zal de zaken in Topdesk rapporteren. Hierbij worden bijvoorbeeld activiteiten/fases vastgelegd en de planning. Tijdens het project worden er in Topdesk aantekeningen gemaakt en bijgehouden hoe het project op bepaalde onderwerpen verloop. Deze rapportages worden door het management en de opdrachtgever beoordeeld. Wanneer een project slecht 'scoort' op zo'n onderwerp grijpt het management in. De maatregelen die hierbij optreden worden toegepast en er wordt weer verder gewerkt aan het project. Wanneer ze in de laatste fase van het project zitten en het project dus afgerond is eindigt het proces. Vooral aan het begin van het proces bij de aanvraag wordt data vastgelegd, maar ook gedurende het project worden zaken vastgelegd per projectfase. Voor de data die tijdens dit proces gerapporteerd wordt moet ik het conceptueel datamodel vastleggen in EA.

Op basis van het BPMN model is er een use case diagram tot stand gekomen.



Figuur 4: Use case diagram

4.2 De data architectuur

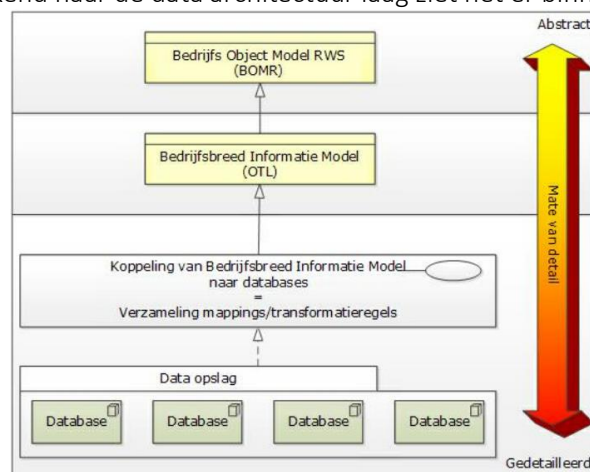
De architectuur ten aanzien van programma's en projecten binnen Rijkswaterstaat is conform TOGAF. TOGAF staat voor The Open Group Architecture Framework en is een standaard voor Enterprise architectuur. In principe kent TOGAF vier domeinen Business architectuur, data architectuur en technische architectuur. Het biedt een language, een aanpak en een reeks aanbevelingen die over alle facetten van Enterprise-architectuur gaan, van organisatie en strategie, tot business en technologie, tot planning en verandermanagement. Dit helpt iedereen het framework beter te begrijpen en zo miscommunicatie te voorkomen (Desfray & Raymond, 2014).

Het doel voor deze architectuur binnen Rijkswaterstaat is om zo structuur en samenhang te geven en dit te bewaken voor de verschillende verzamelingen van gegevens van Rijkswaterstaat. De data is hierdoor actueel, compleet en betrouwbaar en wordt er zo voldaan aan de FAIR principes. Voor verder uitleg van Fair principes zie [hoofdstuk 4.2.1 FAIR principes](#) (Rijkswaterstaat, 2021a).

Als er gekeken wordt naar het totale architectuur landschap dan bestaat dit uit vijf pilaren:

- **Business architectuur:** Hierin wordt de structuur en interactie tussen de organisatie, bedrijfsstrategie, functies, informatiebehoefte en bedrijfsprocessen beschreven.
- **Informatie architectuur:** Hier wordt beschreven hoe de informatiebehoefte het best ingevuld kan worden
- **Data architectuur:** De data architectuur beschrijft de interactie en structuur van de (belangrijkste) data objecten, fysieke data assets, logische data assets en bronnen voor gegevensbeheer.
- **Applicatie architectuur:** Dit beschrijft de structuur en interactie van de applicaties. Het doel hiervan is om de applicaties van Rijkswaterstaat optimaal te laten aansluiten op de informatiebehoefte ten gevolge van de werkprocessen en de informatie behoefte van de partners en klanten van Rijkswaterstaat
- **Technische architectuur:** De structuur en interactie van de platformdiensten technologie componenten worden hierin beschreven. (Rijkswaterstaat, 2021a).

Kijkend naar de data architectuur laag ziet het er binnen Rijkswaterstaat er als volgt uit:



Figuur 5: Samenhang data-architectuur (Rijkswaterstaat, 2021b)

Het begint bij de databases. Rijkswaterstaat geeft de voorkeur aan een informatiemodel dat één op één is overgenomen in de fysieke structuur. Dit is niet altijd mogelijk daarom dat er gebruik moet worden gemaakt van mappings of transformatieregels om zoals nog een juiste vertaalslag te kunnen maken. Het bedrijfsbreed informatiemodel (OTL) is een model meer gedetailleerd in gaat op alle

objecten (die ook in het bedrijfsobjectmodel RWS (BOMR) staan. In deze laag wordt er dieper ingegaan op de objecten. Eigenschappen, composities en relaties met andere objecten komen hierin voor. De door mij gemaakte datamodellen behoren dan ook tot deze laag. In het bedrijfsobjectmodel RWS staan de objecten die nodig zijn bij het proces opgenomen, deze zijn dus ook opgenomen in het OTL. Dit is een hiërarchisch model van de belangrijkste objecten waarvan data over verzameld en versterkt wordt door Rijkswaterstaat. Echter kent het huidige bedrijfsobjecten model RWS wat minpunten. Er is aangegeven dat sommige onderwerpen in dit model veel te gedetailleerd uitgewerkt zijn en andere juist weer onderbreking. Zo wordt er bijvoorbeeld een bedrijfsobject; Netwerkgebruiker gedefinieerd binnen het BOMR. Echter kent Rijkswaterstaat verschillende soorten netwerkgebruikers (bijvoorbeeld Waterinnamebedrijf en Zalm, deze zijn zo onderkend in het waterdomein), maar deze zijn verder niet specifiek terug te vinden op het BOMR. Er moet dus nog een nieuw model komen. Het BOMR uit 2009 dient hiervoor als uitgangspunt. Voor de exacte modellen zie hiervoor [bijlage B: Het bedrijfsobjectmodel RWS](#)

4.2.1 FAIR principes

Zoals eerder aangegeven is het de bedoeling dat data voldoet aan de FAIR principes. Dit zijn een principe binnen Rijkswaterstaat waaraan de data moet voldoen. Er is een kader opgesteld door Margreet Kamp (data architect) om te voorkomen dat voor alle data en databases afzonderlijk bepaald wordt waar het beheer aan moet voldoen. Overal dient volgens dit kader inrichting plaats te vinden, hierin staat dus ook gedefinieerd dat data FAIR moet zijn. Dit kader heeft voornamelijk betrekking op data architecten en mensen die te maken hebben met de data. (Kamp, 2021)

FAIR staat voor vindbaar, toegankelijk, interoperabel en herbruikbaar (findable, accessible, interoperable, reusable). Dit zijn principes waaraan data kan voldoen. De bedoeling hiervan is dat data zo makkelijk te vinden is voor zowel mensen als machines. Data moet dan ook uitgebreide metadata bevatten, zodat het goed te vinden is. De metadata moet een wereldwijd unieke en persistente identifier hebben. Uiteindelijk moet dit ook geregistreerd of geïndexeerd worden in een doorzoekbare bron. Als de data hieraan voldoet dan is het goed vindbaar.

Daarnaast is het ook de bedoeling dat het goed toegankelijk is, zelfs als de data niet meer beschikbaar is zou metadata nog steeds toegankelijk moeten zijn. Het is de bedoeling dat (meta)gegevens kunnen worden opgehaald met behulp van hun identifier door middel van een gestandaardiseerd communicatieprotocol. Dit protocol moet open, gratis en universeel implementeerbaar zijn en moet het authenticatie- en autorisatieprocedure mogelijk maken. Ook zou metadata verwijzingen naar andere (meta)data moeten bevatten om het zo interoperabel te maken. Data zou ook gebruik moeten maken van een formele, toegankelijke en breed toepasbare taal voor kennisrepresentatie.

Met herbruikbaar wordt bedoeld dat data voldoen aan domeinrelevante normen, het duidelijk is waar de data vandaan komen en het duidelijke en toegankelijke gebruikerslicenties bevatten en dat meta(data) zijn uitgebreid beschreven zijn met een vele nauwkeurige en relevante attributen. Wanneer data aan al deze punten voldoet dan is het FAIR data (Wilkinson, 2016).

Persoonlijk heb ik niet veel te maken met de FAIR principes. Ik ben zelf niet verantwoordelijk voor de data zelf, alleen voor de datamodellen. Het enige waar ik echt rekening mee moet houden is dat bijvoorbeeld klassen niet dubbel zijn. De data moet uniek zijn en wanneer er twee dezelfde klassen in een datamodel zijn is de data ook niet meer uniek. Waar ik ook nog mee te maken heb is dat de definities eenduidig zijn. Verder ben ik niet verantwoordelijk voor dat de data bijvoorbeeld een unieke identifier heeft, immers ben ik namelijk niet verantwoordelijk voor de data zelf, maar zoals eerder gezegd alleen van de datamodellen.

5. HET ONDERZOEK

De vraag vanuit de organisatie was om datamodellen ten aanzien van IT-projecten te modelleren en vast te leggen in EA. Tevens was er later ook de vraag hoe de data(modellen) later onderhouden moeten worden. Deze vraag kwam laten vanuit mijn stagebegeleider Rob Smits. Hiervoor wordt er gekeken naar twee frameworks, en uiteindelijk wordt er één aangeraden. Om deze twee vragen te beantwoorden heb ik een kwalitatief onderzoek gedaan.

Het doel hiervan is om door het beantwoorden van de deelvragen uiteindelijk een antwoord te krijgen op de hoofdvraag. De hoofdvraag leidt tot de beroepsproducten. Wanneer op de hoofdvraag een antwoord is gekomen, dan is ook de stage succesvol afgerond.

De hoofdvraag is als volgt:

- Hoe kan ik vanuit de wensen van Rijkswaterstaat nieuwe verbeterde datamodellen maken ten aanzien van IT-Projecten volgens onderbouwde literatuur en hoe kunnen deze datamodellen het best beheert worden?

De hierbij behorende deelvragen zijn:

1. Is EA geschikt voor het modelleren van de gewenste UML-modellen?
2. Wat zijn de requirements waar de modellen aan moeten voldoen?
3. In hoeverre is door het bedrijf dit probleem omtrent modelleren van conceptuele datamodellen al behandeld?
4. Met welke 'regels' moet er rekening worden gehouden bij het modelleren in UML?
5. Wat zegt DAMA DMBOK over het onderhouden/beheren van data(modellen)?
6. Wat zegt De Clouw Data Management Capabilities Framework over het onderhouden/beheren van data(modellen)?

Deelvragen één, twee, drie en vier dragen bij aan het modelleren van de datamodellen. Met de antwoorden die uit deze deelvragen voortkomen kan ik uiteindelijk juiste datamodellen in EA maken. Deelvraag vijf en zes dragen bij aan het advies dat uiteindelijk gegeven wordt. Hierin wordt uitgelegd wat deze frameworks precies inhouden

6. METHODOLOGIE

Tijdens de stage heb ik verschillende onderzoeksmethodes gebruikt. Mijn onderzoek was voornamelijk een kwalitatief onderzoek en dit is ook terug te zien in de soorten onderzoeksmethodes die zijn toegepast. Aangezien het onderzoek ook kijkt naar Rijkswaterstaat als organisatie zelf; hoe gaat de organisatie zelf met bepaalde zaken om? Is er voornamelijk gekeken naar de medewerkers.

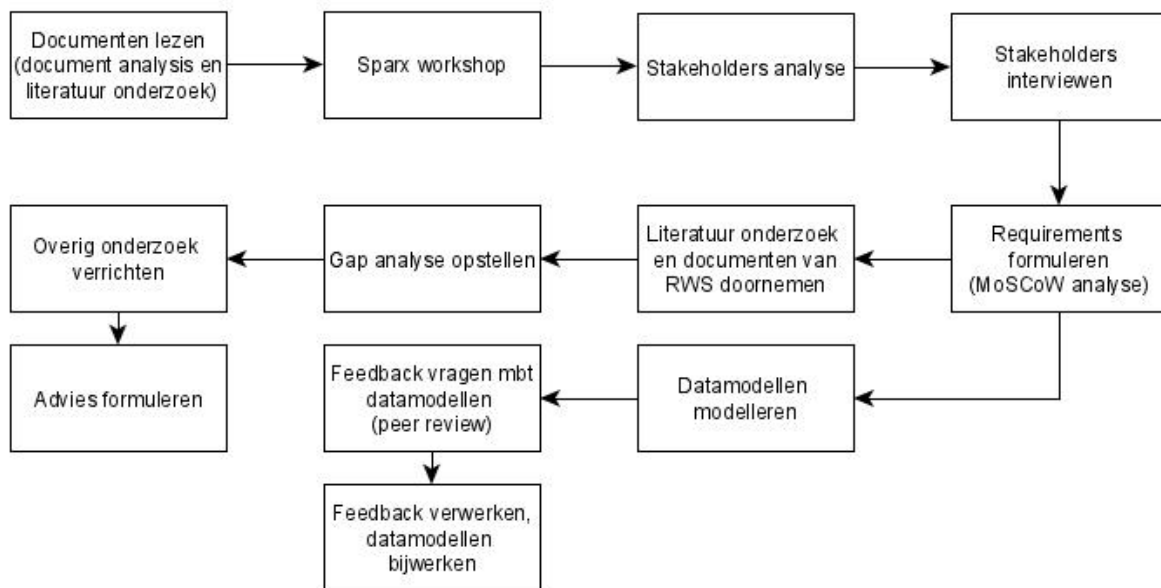
In onderstaande tabel is te zien hoe ik elke onderzoeksmethode heb gebruikt en waar deze is toegepast.

Tabel 3: Overzicht onderzoeksmethodes

<i>Onderzoeksmethode</i>	<i>Gebruikt voor</i>	<i>Toegepast tijdens</i>
Literatuur onderzoek	Verschillende boeken, onderzoeken, etc. zijn geraadpleegd om zo ook te kijken wat de literatuur zegt over onderzochte onderwerpen.	Deelvraag 1, 2, 4, 5, 6,
Interviews	De stakeholders zijn interviewt om zo onder andere de organisatie te analyseren en requirements te achterhalen.	Deelvraag 1, 2, 3, 4
Peer review	Om zo feedback te ontvangen van onder andere de stakeholders.	Datamodellen
Document analysis	Belangrijke informatie binnen de organisatie zelf met betrekking tot de opdracht moet ook bekeken worden. Zo wordt het probleem ook duidelijker en kan er gekeken worden naar een oplossing	Deelvraag 2, 3, Context
MoSCoW analyse	Het in kaart brengen van de requirements waar de datamodellen aan moeten voldoen	Deelvraag 2
Gap analyse	Wordt gebruikt om de huidige situatie te analyseren en na te gaan in hoeverre de verwachting voldoet aan de huidige situatie omtrent conceptuele datamodellen modelleren.	Deelvraag 3
Stakeholders analyse	Deze analyse is gebruikt om alle stakeholders in kaart te brengen en zo na te gaan wie er onder andere geïnterviewd moet worden en hoe belangrijk zij zijn voor het project	Stakeholders
Vragenlijst	Behalve de interviews heb ik ook gebruik gemaakt van een vragenlijst. Sommige stakeholders waren erg druk en gaven de voorkeur aan het invullen van een vragenlijst, in plaats van een interview	Deelvraag 2, 3, 4

Naast dat ik deze onderzoeksmethodes heb toegepast heb ik tijdens het uitvoeren van mijn stage ook de Kanban methode toegepast. De Kanban methode was voor mij de meest voor de hand liggende methode om te gebruiken. Aangezien ik dit project voornamelijk alleen uitvoerde en niet in een team werkte, had ik bijvoorbeeld geen noodzaak aan daily standups. Wel vond ik het voor mij persoonlijk belangrijk om een duidelijk overzicht te hebben van de zaken die nog gedaan moeten worden, gedaan zijn en waar ik nu mee bezig ben. Hiervoor heb ik dan ook een Trello bord opgezet.

Als er gekeken wordt naar de totale stage dan ziet de stage er als volgt uit:



Figuur 6: Schematische weergave aanpak stage

In het begin van de stage heb ik mij vooral ingelezen in de document van Rijkswaterstaat, daarnaast heb ik ook gekeken naar hoe EA werkt. Ik zou ook een workshop hierover krijgen maar omdat ik nog helemaal niks wist van de tool leek hij mij handig om alvast te kijken hoe het werkte. Vervolgens had ik een workshop EA waarbij ik meer leerde over EA en UML. Toen ben ik begonnen met een stakeholders analyse, toen deze in kaart waren gebracht heb ik stakeholders geïnterviewd. Aan de hand van de documenten van Rijkswaterstaat, het literatuur onderzoek en de interviews heb ik de requirements geformuleerd en deze aan de hand van MoSCoW geprioriteerd. Ik nog meer naar de organisatie zelf gaan kijken en ben ik begonnen met modelleren. Toen de modellen af waren heb ik dit aan paar stakeholders (Rutger van Beek, Rob Smits en Arthur de Waard) laten zien. Zij hebben feedback geven op de datamodellen en deze heb ik vervolgens verwerkt. Ook ben ik begonnen aan een gap analyse en uiteindelijk nog meer literatuuronderzoek gedaan om zo tot een advies te komen.

7. ONDERZOEKSRESULTATEN – DEELVRAGEN

Per deelvraag wordt uitgelegd hoe ik dit heb aangepakt en wat uiteindelijk de resultaten zijn. Vervolgens leidt dit tot de hoofdvraag en komt hier ook een antwoord op.

7.1 Is EA geschikt voor het modelleren van de gewenste UML-modellen?

EA is de tooling die binnen Rijkswaterstaat gebruikt wordt. Voor dat Rijkswaterstaat EA in gebruik heeft genomen, gebruikte de organisatie Troux Architect en Visio. Wegens onder andere de kosten is Rijkswaterstaat uiteindelijk overgestapt naar EA. Deze was in eerste instantie ook voor de architectuureen ander doeleind aangeschaft, maar omdat het platform zo uitgebreid is, is er uiteindelijk ook voor gekozen om de datamodellen ook in EA te modelleren.

Aanpak

Om deze vraag te kunnen beantwoorden wordt eerste gekeken naar EA is. Wat kan je meedoen? Vervolgens wordt er gekeken naar waarom Rijkswaterstaat hiervoor heeft gekozen.

Resultaat

EA is een tool voor visualiseren en modelleren van business en IT-systemen en kan gebruikt worden voor software en systeem engineering. Het is prima geschikt voor het modelleren van database en business processen, system engineering en simulatie, projectmanagement en debuggen. Het heeft ingebouwde mogelijkheden voor requirements management dat helpt om met behulp van UML, SysML, BPMN, etc. Specificaties specificaties op hoog niveau te traceren om zo modellen te analyseren, ontwerpen, implementeren, testen en onderhouden. Verder helpt het ook bij het bouwen van onderhoudbare systemen en door gebruikt te maken van ingebouwde rapportages en documentje. EA heeft een high performace waar het snel grote modellen kan laden. Door de grote model repository maakt de tooling het mogelijk om (grote) teams aan hetzelfde project te laten werken. Naast dat dit samenwerken aan gedeelde project makkelijk maakt, beschikt de tooling ook nog een geïntegreerde versiebeheermogelijkheden en inzetbare Cloud gebaseerde servers (Sparx Systems, 2022a).

Sparx biedt ook de mogelijkheid aan om modellen te simuleren. Door gebruik te maken van triggers zoals het indrukken van een knop of schakelaar, of het ontvangen van een bericht om de uitvoering van de simulatie te besturen. Zo kan er controleert worden of modellen kloppen en kunnen bedrijfssystemen beter begrepen worden. De simulatie kan je zelf helemaal regelen. Het kan helemaal gemanipuleerd worden, zo ook variabelen en de snelheid van simulatie waardoor mogelijk 'onzichtbare' stappen dan beter de analyseren zijn. Ook heeft het built-in requirement managementfuncties waardoor het ook hiervoor gebruikt kan worden. Het kan onder andere gebruikt worden voor het definiëren van requirement modellen, het zoeken en rapporteren over requirements en een impactanalyse uitvoeren op basis van voorgestelde veranderen van requirements. EA is ook handig voor documentatie. De tooling kan documenten genereren met informatie over bijvoorbeeld een model en daarnaast ondersteund het ook verschillende talen waaronder: C, C++, C#, Java, python, PHP, etc. Door de built-in source code editor is het mogelijk om van model naar course code te gaan (Sparx Systems, 2022a).

UML modelleren is ook prima te doen. Er zijn verschillende manieren om bedrijfsprocessen te modelleren. Hierbij wordt UML als basis modelleertaal gebruikt. De fundamente van EA Enterprise Architect zijn gebouwd op de UML 2. Het vult UML 2.5 aan met BPMN-ondersteuning en

uitbreidingselementen voor analyse, vereistenbeheer en procesbeheer (zoals wijzigings-, functie- en probleemelementen). Het stelt de gebruiker in staat om bedrijfsprocessen (in UML-format) te visualiseren (Sparx Systems, 2022a).

Tevens is dit ook handig voor projectmanagers. Zo is er de mogelijkheid om resources aan elementen toe te wijzen, een inschatting te maken van risico's en inspanning, het inschatten van de grootte van een project en de complexiteit ervan. Met het programma kan je ook Kanban toepassen en Gantt tabellen maken (Sparx Systems, 2022a).

Voor het totaal overzicht van functionaliteiten binnen EA, zie [bijlage E: Functionaliteiten per editie Sparx Enterprise Architect](#)

Het is dus een veelzijdige tooling en is marktleider op gebied van architectuur tools. Dit is ook een van de redenen waarom Rijkswaterstaat gekozen heeft voor EA. Doordat ze hebben gekozen voor de Ultimate edition kunnen ze het meest uit de tool halen, waardoor er ook veel mee gedaan kan worden. Een andere reden waarom er gekozen is EA (en ook de hoofdreden) was de prijs. Rijkswaterstaat vond Troux te duur worden en is daarom opzoek gegaan naar een alternatief.

Als er puur gekeken wordt naar functionaliteiten binnen Sparx Enterprise Architect, dan lijkt het alsof Sparx meer aan te bieden heeft. Kijkend naar de functionaliteiten van Troux dan lijkt het erop dat Troux meer bedoeld is voor projectmanagement. Het heeft veel verschillende functionaliteiten ter bevordering van de samenwerking. Zo kunnen Zoom meetings vanuit de projectplace gedaan worden (Planview projectplace, 2021). Echter heb ik later te horen gekregen dat Troux is overgenomen door planview. Rijkswaterstaat heeft echt de versie voor de overname, waardoor deze functionaliteiten nu niet aanwezig zijn.

Het ziet er niet uit alsof EA veel mist en is ook prima geschikt om datamodellen te modelleren. Het enige nadeel wat eventueel kan zijn van deze tooling is dat het in het begin nog lastig kan zijn om het helemaal te begrijpen. Het is zo'n brede tooling en het biedt zoveel mogelijkheden dat het te veel kan worden. Dit heb ik zelf persoonlijk ook gemerkt, maar ook binnen de organisatie kwam ik mensen tegen die het ook ingewikkeld vonden.

7.2 Wat zijn de requirements waar de modellen aan moeten voldoen?

Aanpak

Voor het achterhalen van de requirements heb ik verschillende medewerkers bij Rijkswaterstaat geïnterviewd. Voor een transcript van de interviews, waarin de gestelde vragen te lezen zijn [zie bijlage C: interviews](#). Hierbij heb ik Rob Smits (informatie manager), Margreet Kamp (data architect), Peter Bernhard (senior adviseur) en Victor van der Kloet (informatieanalist) geïnterviewd. Deze mensen zijn nauw bij de opdracht betrokken of hebben ervaring met modelleren. Daarnaast heb ik ook van een geïnterviewde een document gekregen waarin de algemene requirements van data architectuur in staan, hier stonden echter niet veel requirements in met betrekking tot datamodellen. Omdat er uit de interviews niet veel specifieke requirements met betrekking tot de datamodellen kwamen heb ik ook gekeken naar datamodellen van andere mensen binnen Rijkswaterstaat hoe hen datamodellen eruit zien om te kijken of er overeenkomsten zijn op gebied van modelleren. Er is ook rekening gehouden met UML requirements (Bellekens, 2018).

Om alle requirements te analyseren en te prioriteren is er gebruik gemaakt van de MoSCoW methode. MoSCoW staat voor:

- **Must have:** Requirements geprioriteerd onder dit label zijn harde eisen. Hier moet hoe dan ook aan voldaan worden. Als niet aan deze requirements wordt voldaan dan is het beoogde resultaat niet behaald.
- **Should have:** Hieronder vallen de requirements die een hoge prioriteit bevatten maar niet perse verplicht zijn.
- **Could have:** Wanneer er tijd over is, zal er gekeken worden naar deze requirements. Het is niet noodzakelijk en kan meer gezien worden als een wens in plaats van een harde eis.
- **Won't have:** Dit zijn requirements waar gedurende de stage niet aan gewerkt worden. Deze requirements hebben geen toegevoegde waarde en vallen ook niet binnen de scope.

Deze methode laat dus zien aan welke requirements voldaan moet en hoe belangrijk een requirement voor het product is.

Resultaat

De analyse ziet er als volgt uit:

Tabel 4: MoSCoW analyse

Prioriteit	Requirement
M	Inhoud van Microsoft Word documenten op zijn genomen in het model
M	Definities van klassen en attributen staan in het model
M	Relaties hebben een duidelijke beschrijving
M	Het model is makkelijk te onderhouden/managen
M	Relaties tussen klassen zijn correct
M	Datamodellen staan in EA
M	Datamodel staat in UML formaat
S	Er wordt onderscheid gemaakt tussen verschillende soorten relaties
S	De relatielijnen zijn horizontaal of verticaal
S	Klassen beginnen met een hoofdletter
C	Datamodellen correspondeert met andere al bestaande datamodellen in de architectuur
C	Klassen hebben waar kan dezelfde grootte
C	Generalisatie relaties wijzen omhoog of horizontaal
C	Klassen staan op dezelfde hoogte
C	Niet te veel relaties moeten elkaar kruisen
C	Attributen beginnen met een kleine letter
W	Klassen binnen het datamodel bevatten operations

Om de prioritering te bepalen is er gekeken naar wat direct de eisen waren van de stakeholders. Deze eisen zijn gelabeld als must-have. Requirements die naar voren kwamen door te kijken hoe andere mensen binnen Rijkswaterstaat hun datamodellen modelleerden zijn gelabeld als should-have. Ook requirements die volgens externe bronnen gehanteerd moeten worden en ook zijn toegepast door medewerkers binnen Rijkswaterstaat zelf vallen hieronder. Deze requirements vallen niet onder must-have omdat deze namelijk niet uitdrukkelijk benoemd waren door de stakeholders. Eisen waaraan een datamodel die volgens best practice aan zou moeten voldoen zijn gelabeld als could-have (Bellekens, 2018). Deze requirements zijn namelijk ook niet altijd door Rijkswaterstaat gehanteerd en verder ook niet benoemd door de stakeholders. Dit zijn dus eisen waar het datamodel aan moet voldoen volgens externe bronnen. Eisen gelabeld als won't-have zijn requirements waar een datamodel volgens externe bronnen aan moet voldoen maar niet relevant zijn voor de organisatie en ook niet door andere medewerkers binnen Rijkswaterstaat zijn toegepast. Hier heb ik ook met de stakeholder over gehad en er is naar voren gekomen dat deze requirements ook niet nodig zijn binnen het model.

7.3 In hoeverre is door het bedrijf dit probleem omtrent modelleren van conceptuele datamodellen al behandeld?

Aanpak

Om deze deelvraag te onderzoeken heb ik voornamelijk gekeken naar het bedrijf zelf. Ik heb zowel gewone gesprekken gevoerd met stakeholders, (kennismakingsgesprekken en gesprekken ten aanzien van kennisuitwisseling) als interviews met stakeholders, waarbij ik specifieke vragen stelden. Hierbij heb ik Rob Smits, Victor van der Kloet, Margreet Kampen en Peter Bernhard heb geïnterviewd. Door hen betrokkenheid/functie konden zij een goed vormen van wat er nu gaande was met betrekking tot het modelleren van conceptuele datamodellen binnen de organisatie. Daarnaast heb ik ook gekeken naar documenten binnen Rijkswaterstaat zelf om zo te kijken hoe dit er allemaal voor staat. Middels de SERVQUAL-methode heb ik gekeken naar de verwachting en de werkelijkheid. Hier ben ik ingegaan op drie dimensies die het meest relevant waren. De overige drie dimensies waren niet heel relevant. Ik heb ervoor gekozen om deze methode toe te passen omdat hierbij wordt na gegaan wat de verwachting was en wat uiteindelijk de realiteit is.

Het IstSoll model is ook een vorm van een gap-analyse maar hierbij zag ik voornamelijk dat gedaan werd aan de hand van prestatie indicatoren, dus bijvoorbeeld de productie of de efficiency van een geleverde output (output producten vs aantal mensen). Dit is in deze situatie moeilijk na te gaan want hier wordt niet gesproken in absolute of relatieve waarden. Een andere analyse voor het meten van een gap was de SWOT analyse, maar deze focust weer meer op de concurrent van de organisatie. Door de eigen zwaktes en sterktes in kaart te brengen en te kijken naar interne en externe bedreigingen kan zo gekeken worden naar mogelijke kansen voor de organisatie en hoe de organisatie zich vervolgens kan onderscheiden van de concurrentie. Dit leek mij moeilijk toe te passen op deze situatie omdat de focus op een ander punt ligt namelijk de huidige situatie analyseren en kijken hoe ver het al is met het modelleren van datamodellen.

Uiteindelijk ben ik tot de conclusie gekomen dat een gap-analyse middels de SERVQUAL-methode de beste manier is om na te gaan in hoeverre Rijkswaterstaat al bezig is met het modelleren van conceptuele datamodellen.

De SERVQUAL-methode kijkt naar de verwachten en uiteindelijk de werkelijkheid. Aan de hand van scores vergelijkt het dan hoe de stand van zaken ervoor staat en kan er zo een beeld komen van de gaps. Er zijn vijf dimensies:

- Betrouwbaarheid: Hoe accuraat en eerlijk is het?
- Responsiviteit: Hoe snel wordt er gereageerd?
- Zekerheid: kan je zeker zijn van de service en daarmee verzekerd zijn van de juiste informatie?
- Empathie: Hoe erg speelt dit in op de wensen en behoeften?
- Uitstraling: Hoe komt het over?

Hierbij heb ik gekeken naar hoe belangrijk zo'n dimensie was voor de analyse. Vervolgens ben ik nagegaan wat het streven is, dus bijvoorbeeld:

Bij dit bedrijf zullen visueel weergegeven datamodellen aanwezig zijn. Dit is namelijk het ideaal, schematische weergave van de datamodellen, volgens heb ik mijn verwachting en de waarneming die ik heb gedaan op dit punt een score gegeven. Deze score wordt vervolgens toegelicht.

Resultaat

Bij het uitvoeren van de analyse heb ik gekeken naar drie dimensies: Betrouwbaarheid, zekerheid en uitstraling. Hieronder staat per dimensie de verwachting en de waarneming geformuleerd. Op basis

van de 5-Punts Likert schaal wordt er een score aan gekoppeld. 1 = Helemaal oneens, 2 = een beetje oneens, 3 = Neutraal, 4 = een beetje eens, 5 = helemaal mee eens.

Bij verwachting wordt dus uitgedrukt in hoeverre dit verwacht werd. Bij werkelijkheid wordt aangegeven in hoeverre het aan de eerdere verwachting doet. Hieronder is per dimensie aangegeven wat de verwachting en de werkelijkheid was. Voor elke verwachting/waarneming staat een cijfer. Tijdens het toelichten van de score wordt naar dit cijfer gerefereerd.

Tabel 5: Gap-analyse uitstraling

Dimensie: Uitstaling**Scores per vraag: 5-Punts Likert schaal: 1 (helemaal oneens) – 5 (helemaal mee eens)**

Verwachting (V)	Score	Werkelijkheid/waarneming (W)	Score
V1. Bij dit bedrijf zullen visueel weergegeven datamodellen aanwezig zijn	2	W1. In dit bedrijf zijn visueel weergegeven datamodellen aanwezig	2
V2. Bij dit bedrijf zullen er standaarden zijn voor modelleren	4	W2. Bij dit bedrijf zijn er standaarden voor modelleren	2
V3. Bij dit bedrijf zal er één modelleertool aanwezig zijn waar iedereen gebruik van maakt.	3	W3. Bij dit bedrijf is er één modelleertool aanwezig waar iedereen gebruik van maakt.	3
V4. Er zal uniformiteit zijn binnen de datamodellen van de organisatie.	4	W4. Er is uniformiteit binnen de datamodellen van de organisatie.	3

Toelichting op de score: In eerste instantie werd er niet heel erg verwacht dat binnen Rijkswaterstaat al veel datamodellen visueel weer waren gegeven. Aangezien de opdracht hier overging was de verwachting dat het al minimaal was. De werkelijkheid was dus ongeveer hetzelfde als de verwachting. Er was niet heel veel op gebied van datamodellen. De verwachting bij nummer 2 was dat Rijkswaterstaat meer standaarden had voor modelleren. De aanname was dat Rijkswaterstaat een groot bedrijf is en het daarom vast wel zou werken met standaarden of werkwijzers. Dit doet Rijkswaterstaat ook wel maar nog niet op gebied van conceptuele datamodellen. Echter bleek het zo dat die vanuit het bedrijf zelf er niet tot nauwelijks waren. Ze gebruiken wel frameworks en hebben voor bepaalde processen ook een werkwijze maar voor datamodelleren in UML hebben ze dit nog niet. Ze kijken wel naar de meer algemene standaarden, maar Rijkswaterstaat zelf heeft geen eigen standaard voor het modelleren van datamodellen. De werkelijkheid verschilde dus van de verwachting.

Verwachting 3 was gelijk aan de waarneming. De verwachting was dat juist omdat het een grote organisatie is, dat mensen ook wel verschillende programma's gebruiken. In de werkelijkheid is dit ook half zo. Momenteel proberen ze alles met EA te doen, maar er staan nog modellen in Troux. Deze zijn nog niet overgezet. Ze proberen wel alles in één tooling te doen, maar het staat nog wel in verschillende toolings.

Kijkend naar de datamodellen was er wel een redelijk hoge verwachting dat er uniformiteit was binnen de datamodellen. Dit is ook wel ongeveer zo. UML-regels worden aangehouden, maar niet elk model wordt op dezelfde manier gemodelleerd. Het ene model is zeer geordend en het andere model iets minder. Dit is voornamelijk afhankelijk van de persoon die eraan heeft gewerkt, of in welke fase het model zich bevindt. De werkelijkheid was dus een klein beetje anders dan de verwachting.

Tabel 6: Gap-analyse betrouwbaarheid

Dimensie: Betrouwbaarheid**Scores per vraag: 5-Punts Likert schaal: 1 (helemaal oneens) – 5 (helemaal mee eens)**

Verwachting (V)	Score	Werkelijkheid/waarneming (W)	Score
-----------------	-------	------------------------------	-------

V5. Bij dit bedrijf zullen de Microsoft Word document volledig en betrouwbaar zijn	5	W5. Bij dit bedrijf zijn de Microsoft Word document volledig en betrouwbaar	3
V6. Bij dit bedrijf zullen datamodellen goed omschreven zijn.	4	W6. Bij dit bedrijf zijn datamodellen goed omschreven.	4
V7. Bij dit bedrijf zal het duidelijk zijn wat er met bepaalde begrippen binnen de datamodellen bedoeld wordt.	4	W7. Bij dit bedrijf is het duidelijk wat er met bepaalde begrippen binnen de datamodellen bedoeld wordt.	4

Toelichting op de score: De datamodellen zijn sterk afhankelijk van informatie in de Microsoft Word documenten, de verwachting was dus ook dat deze documenten volledig en betrouwbaar waren. In de werkelijkheid zijn deze documenten nog niet volledig. Het is een levend document en wordt nog steeds veranderd.

Verwachting zes, dat datamodellen goed omschreven waren, had te maken met de omschrijving van de opdracht. De datamodellen moesten onder andere definitieproblemen voorkomen. In de werkelijkheid blijkt het ook zo dat elk model, klasse en relevante attributen en omschrijving bevatte. De verwachting kwam dus ook overeen met de werkelijkheid. Verwachting zeven sluit hier dan ook goed bij aan. Dat de datamodellen goed omschreven zijn zorgt er dus ook voor dat het duidelijk was wat er met bepaalde begrippen bedoeld wordt.

Tabel 7: Gap-analyse zekerheid

Dimensie: Zekerheid				
Scores per vraag: 5-Punts Likert schaal: 1 (helemaal oneens) – 5 (helemaal mee eens)				
	Verwachting (V)	Score	Werkelijkheid/waarneming (W)	Score
V8. De standaarden voor modelleren zullen bij de betrokken bekend zijn.	4		W8. De standaarden voor modelleren zijn bij de betrokken bekend	2
V9. Bij dit bedrijf zal het duidelijk wie verantwoordelijk is voor wat ontremd data modelleren.	3		W9. Bij dit bedrijf is het duidelijk wie verantwoordelijk is voor wat ontremd data modelleren	4

Toelichting op de score: Verwachting 8 had een score van 4 omdat er gedacht werd dat zoals eerder gezegd standaarden binnen Rijkswaterstaat met betrekking tot datamodellen modelleren aanwezig waren. Dit bleek echter niet het geval te zijn, daardoor wordt het ook voor de betrokkenen lastig om kennis te hebben over standaarden. Deze zijn er immers niet echt aanwezig binnen de organisatie.

Er waren niet echt veel verwachtingen over nummer 9. Datamodelleren zoals ze het nu doen is nog vrij nieuw voor de organisatie, daarom dat werd verwacht dat er nog niet echt duidelijk een persoon hiervoor was aangewezen. In de werkelijkheid is er officieel gezien ook niet iemand aangewezen die hierover gaat. Echter hebben mensen toch (onbewust) deze taak op zich genomen waardoor het nu wel duidelijk is wie hier over gaat. Daarnaast wordt er is sommige modellen aangegeven wie hier verantwoordelijk voor is en daarnaast kan je binnen EA zien wie het model heeft aangemaakt. Zo is wel te zien wie over een bepaald model gaat.

De conclusie die ik hieruit kan trekken is dat er nog niet heel veel is op gebied van datamodellen. Het streven is dat elke stelling een score heeft van 5. Dit is momenteel nog niet het geval. Vooral in de dimensie uitstraling is er nog ruimte voor verbetering. Er is namelijk nog niet heel veel uniformiteit binnen de manier waarop datamodellen tot stand komen. Wel zijn er al Microsoft Word documenten. Het zijn nog levende bestanden (ze worden nog bijgewerkt), maar er zijn wel al duidelijke definities geformuleerd. Kijkend naar dimensie zekerheid is er vooral te zien dat het duidelijk is wie waarvoor verantwoordelijk is, maar zoals eerder gezegd zijn er gewoon geen standaarden met betrekking tot modeleren

7.4 Met welke 'regels' moet er rekening worden gehouden bij het modelleren van een UML Klassendiagram?

Aanpak

Voor het beantwoorden van deze vraag heb ik naar drie dingen voornamelijk gekeken:

1. Wat heeft de literatuur te zeggen over UML, hiervoor kijk ik onder andere naar het boek: *Unified Modeling Language User Guide* geschreven door G. Booch en J. Rumbaugh.
2. Welke regels zijn belangrijk voor Rijkswaterstaat zelf?
3. De cursus UML en EA die van Rijkswaterstaat heb ontvangen. Deze is gegeven door the collective the systems integration company. Hier heb ik ook rekening mee gehouden.

Resultaat/uitleg

Volgens de opdracht is het de bedoeling dat er een model tot stand komt met betrekking tot IV-projecten. Dit moest een UML-klassendiagram zijn. UML kent natuurlijk meer soorten modellen, maar om het binnen de scope te houden wordt hier alleen maar gekeken naar 'regels' binnen UML-klassendiagram. Ik ben voornamelijk ingegaan op klassen en relaties. Tijdens de Workshop EA ging het daar ook om. Andere zaken zoals interface en packages zijn niet aan de orde gekomen. Ook als ik keek naar andere UML datamodellen

Een klassendiagram is een statische weergave van een systeem. Het beschrijft als het ware de structuur van een systeem. Het laat de klasse met hen attributen en operations en bijhorende relaties zien. Een klassendiagram kent dus verschillende onderdelen. Deze onderdelen worden hieronder verder toegelicht.

7.4.1 Klasse

Een klasse is een beschrijving van een verzameling objecten die dezelfde eigenschappen, operations en relaties bevatten. Elke klasse heeft zijn eigen naam wat zich onderscheidt van een andere klasse. Een voorbeeld van een klasse kan zijn: Medewerker. Er werken verschillende mensen (objecten) in een bedrijf maar ze hebben allemaal dezelfde eigenschappen. Bijvoorbeeld, elke medewerker heeft een medewerkers ID, een naam, een afdeling, etc. Dus Piet heeft medewerkers ID 111 en behoort tot de afdeling human resource, terwijl Gerrit medewerkers ID 222 heeft en tot de afdeling IT behoort. Piet en Gerrit zijn objecten en zijn exemplaren binnen de klasse Medewerker.

Deze eigenschappen (medewerkers ID, een naam, een afdeling, etc.) worden ook wel attributen genoemd. Een attribuut is dus een eigenschap van een klasse. Een klasse kan ook meerdere attributen hebben of zelfs geen. Een attribuut bestaat uit een naam en een datatype. De waarde van een attribuut kan verschillende soorten datatypes zijn. Echter kan een attribuutwaarde niet verwijzen naar een andere klasse. In plaats van de klasse op te nemen als attribuutwaarden zal er een relatie tussen de twee klassen komen. Wat een relatie is wordt uitgelegd in [hoofdstuk 7.4.2: Relaties](#). Een attribuut begint met een kleine letter. (IBM, 2015)

Naast dat een klasse dus een naam en attributen bevat, kan het ook nog operations/methodes hebben. Een operation is een abstractie van iets wat je kan doen met een object. Dit wordt ook gedeeld door alle objecten van die klasse. Er wordt hier verder niet op ingegaan omdat deze niet relevant zijn voor de opdracht. Operations komen niet in de datamodellen voor omdat het puur gaat om de definities van objecten met hen attributen en hen relaties.

Een attribuut en een operation hebben ook een visibility (zichtbaarheid). Dit is een eigenschap van een attribuut/operation dat laat zien of andere klassen deze kunnen gebruiken. Er zijn drie verschillende visibilities:

- Public: Alle objecten kunnen gebruik maken van de attribuut/operation. Dit wordt aangeduid met een '+'
- Protected: alleen het object zelf of subklassen hiervan kunnen de attribuut/operation gebruiken. Dit wordt aangeduid met een '#'
- Private: Alleen het object zelf kan deze attribuut/operation gebruiken. Dit wordt aangeduid met een '-'

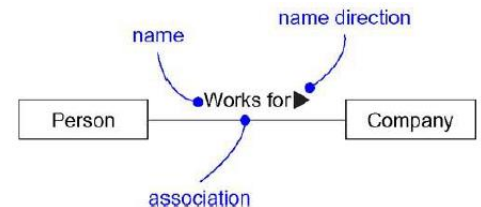
Daarnaast bestaat er ook zoiets als een interface. Een interface is een verzameling van operations die worden gebruikt om het gedrag van een klasse of een component te specificeren. Een interface beschrijft dus het zichtbare gedrag van een element. Het kan het complete gedrag of een gedeelte van het gedrag van een klasse vertegenwoordigen.

7.4.2 Relaties

Wanneer je een model hebt met klassen is al snel te zien dat er maar weinig klassen zijn die op zichzelf staan. De meeste hebben wel te maken met andere klasse. Wanneer een klasse te maken heeft met een andere klasse dan wordt er gesproken van een relatie. Er zijn verschillende soorten relaties tussen klassen.

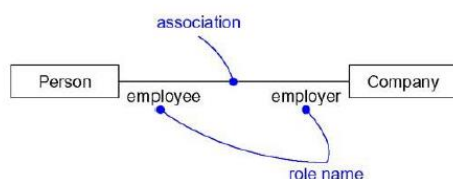
Association: Een association laat zien dat elementen met elkaar verbonden zijn. Dit kan eenzijdig of tweezijdig zijn. Tweezijdig houdt in dat informatie van beide kanten beschikbaar is. Stel je hebt een klant en een bestelling. Tweezijdige association houdt in dat vanuit de bestelling informatie van de klant opgehaald kan worden en vanuit de klant informatie van de bestelling opgehaald kan worden. Eenzijdig houdt in dat er alleen maar vanuit één element informatie van het andere element kan opgehaald worden, maar dit geldt niet andersom. Dus bijvoorbeeld, vanuit de kant kan je informatie krijgen over de bestelling, maar niet vanuit de bestelling over de klant.

Een association kent ook verschillende onderdelen. Het kan dus een richting hebben (eenzijdig of tweezijdige association), het heeft een naam, rollen en multiplicity (verder in het document staat dit uitgelegd). Een association heeft dus een naam. De naam wordt gebruikt om de relatie te omschrijven. Hierdoor wordt dus duidelijk gemaakt wat er met deze relatie wordt bedoelt. Bijvoorbeeld persoon werkt voor bedrijf. Zie in figuur 6 een voorbeeld. (Booch & Rumbaugh, 2005)



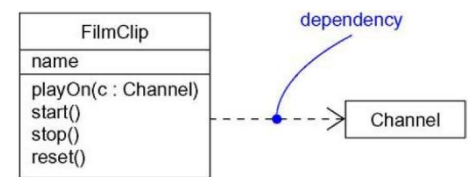
Figuur 7: Voorbeeld association namen (Booch & Rumbaugh, 2005a)

Daarnaast kunnen er ook rollen omschreven worden. Als er een associatie is dan zijn er verschillende rollen bij betrokken. In onderstaande afbeelding figuur 7, zijn twee klassen geïllustreerd. Person en Company, hierbij heeft de klasse Person de rol employee. De klasse Company heeft hierbij de rol employer (Loonen, 2010) (Booch & Rumbaugh, 2005)



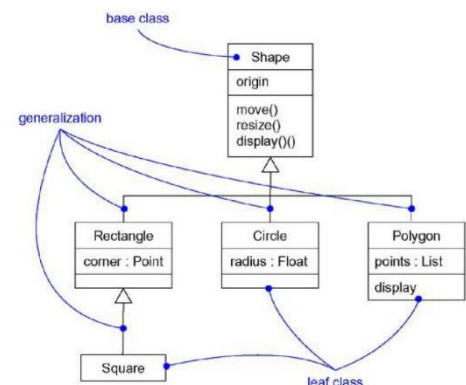
Figuur 8: Voorbeeld association rollen (Booch & Rumbaugh, 2005b)

Dependency: Dit is een directe relatie dat laat zien dat één of meerdere klassen/elementen een ander element nodig heeft/gebruikt voor specificatie of implementatie. ZO kan het dus zijn dat een klasse een andere klasse gebruikt als argument voor een operation. Wanneer dus de gebruikte klasse verandert, kan ook de werking van de andere klasse beïnvloed worden. Zulke relaties hebben meestal geen naam en worden weergegeven middels een stippellijn.



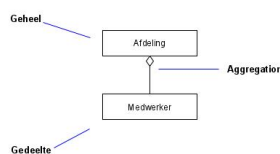
Figuur 9: Voorbeeld dependency relatie (Booch & Rumbaugh, 2005c)

Generalisation/inheritance: Dit is een relatie tussen een algemeen element (hoofdgroep), wordt ook wel een superclass of een parent class genoemd, en een meer gespecificeerd element. Dit wordt ook wel een sub class of child class genoemd. Dit betekent dat een child class de attributen en operations van een parent class overerft maar andersom geldt dit niet. In principe zou je dus de parent class kunnen vervangen voor een child class. Echter moeten de child classes wel van elkaar verschillen en van de parent class. Dus elke child class moet zijn eigen operations en/of attributen hebben. Verschil moet erin zitten anders valt het niet te onderscheiden van elkaar. Een voorbeeld hiervan is bijvoorbeeld parent class: Vorm (superclass), echter kennen we meerdere soorten vormen. Denk aan rechthoek, cirkel of hexagoon. Dit kunnen child classes zijn. Zoals in figuur 7 te zien is, is dat de parent class één attribuut heeft en drie operations. Deze erven de child classes dus over, maar zelf hebben ze ook hen eigen attributen/operations. Hierdoor is elke child class anders. Wanneer een klasse geen parent class heeft maar één of meerdere child classes wordt dit ook wel een rootclass genoemd. Als een klasse zelf geen child classes heeft wordt dit een leaf class genoemd. In figuur 7 is ook te zien dat dit gaat om single inheritance, wanneer een klasse meerdere parent classes heeft dan wordt er gesproken over multiple inheritance. In het model ziet dit eruit als een lijn met een gesloten pijl. Ze hebben meestal geen naam.



Figuur 10: Voorbeeld Generalisation relatie (Booch & Rumbaugh, 2005d)

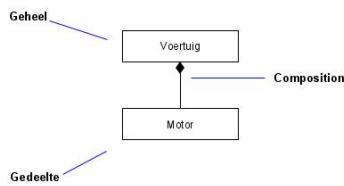
Aggregation: een aggregation is een speciaal type association. Maar waar een association laat zien dat klasse op hetzelfde niveau zit (de een is niet belangrijker dan de ander) werkt dat bij een aggregation iets anders. Sommige dingen bestaan uit/bevatten meerdere kleinere dingen. Je hebt een geheel en dat bestaat uit meerdere delen. Denk bijvoorbeeld aan medewerker en afdeling. Een afdeling bestaat uit meerdere medewerkers. Dit is een aggregation. Belangrijk om hierbij te onthouden: stel een van de elementen valt weg, dan kan de andere nog wel voortbestaan. Een ander element heeft dus geen invloed op het ontstaan of voortbestaan van het element. Een aggregation kan gezien worden als 'bevat', 'bestaat uit' of 'hoort bij' relatie. Het wordt afgebeeld als een witte diamant. (Loonen, 2010)



Figuur 11: Voorbeeld Aggregation relatie

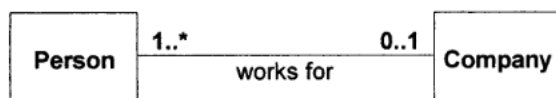
Composition: Een composition lijkt op een aggregation. Het is ook een soort van aggregation. Het grote verschil zit echter in de afhankelijkheid. Bij een aggregation is levensbestaan van het ene element niet afhankelijk van het andere element. Bij Composition is dat wel zo. Ook bij deze relatie kan het gezien worden als een geheel dat bestaat uit meerdere onderdelen, alleen wanneer het geheel wegvalt, vallen ook alle onderdelen weg. Voorbeeld hiervan is als je een klasse voertuig en een klasse motor hebt. Wanneer het voertuig wegvalt, verdwijnt de motor ook. En als er een voertuig

ontstaat, dan ontstaat er ook automatisch een motor. Het voortbestaan van de motor is afhankelijk van het voertuig. Deze relatie wordt afgebeeld middels een zwarte diamant (Loonen, 2010) .



Figuur 12: Voorbeeld composition relatie

Multiplicity: multiplicity geeft aan hoeveel objecten er deelnemen in een relatie. Bijvoorbeeld: één persoon werkt voor nul of één bedrijf, maar bedrijf heeft één tot meerdere werknemers. Dan ziet dat in een datamodel er als volgt uit:



Figuur 13: Voorbeeld multiplicity (G'Enova et al., 2002, Figuur 1)

Er zijn dus verschillende waardes. De volgende waarde zijn er:

Tabel 8: Multiplicity

Mogelijke waardes	Betekenis
1:1	Precies één object
0..*	Nul of meer objecten
1..*	Ten minste één object, maar kan ook meer zijn
m:m	Precies <i>m</i> objecten. Dit kan zijn 5:5, dus precies 5 objecten
m:n	Ten minste <i>m</i> objecten, maar niet meer dan <i>n</i> objecten. Het moet dus tussen <i>m</i> en <i>n</i> in zitten.

7.5 Wat zegt DAMA DMBOK over het onderhouden/beheren van data(modellen) ?

Aanpak

Voor het beantwoorden van deze deelvraag heb ik zowel gebruik gemaakt van literatuur onderzoek als interviews afgenomen van medewerkers binnen Rijkswaterstaat. Hierbij heb ik eerste gekeken naar relevante kennisgebieden voor Rijkswaterstaat. Dit heb ik gedaan door te kijken wat medewerkers binnen Rijkswaterstaat belangrijk vonden. Volgens heb ik gekeken naar wat DAMA DMBOK hierover zegt en de gekozen kennisgebieden uitgelegd.

Ik heb naar dit framework gekozen om drie redenen:

- De organisatie gaf aan dat ze hier benieuwd naar waren, ze wilde hier graag meer over weten.
- Dit werd ook al in de documentatie van opdracht benoemd, dat het handig was om naar te kijken
- DAMA DMBOK is het bekendste data management framework,

Resultaat/uitleg

Om het hele framework uit te werken kost te veel tijd en niet alles heeft betrekking op de scope, daarom dat er gekeken is naar 2 kennisgebieden: Data governance en Data architecture. Na gesprekken met de stakeholders en in het achterhoofd houdend dat de opdracht voornamelijk gaat

over conceptuele datamodellen en onderhoud van deze modellen is er tot de conclusie gekomen dat deze twee functie gebieden het meest belangrijk zijn en binnen de scope vallen.

DAMA DMBOK framework wordt gezien als een van de bekendste frameworks op gebied van datamanagement. Dit boek gaat over 11 kennisgebieden waarbij het draait om databeheer (SynTouch, 2020) (Bluefrog B.V., 2021) (Dama International et al., 2017). Het kent de volgende 11 kennisgebieden: Data governance, Data architecture, Data modeling & Design, Data Storage & Operations, Data security, Data Integration & Interoperability, Documents & Content, Reference & Master Data, Data Warehousing & Business Intelligence, Metadata, Data Quality.

Data governance: Elk bedrijf heeft wel data, en maakt aan de hand deze data ook beslissing. Hierbij is alleen de vraag of zij ook een goede data governance hebben. Organisaties die dit namelijk wel hebben kunnen de waarde van hen data vergroten.

Binnen DAMA DMBOK wordt data governance als volgt gedefinieerd: Het is het uitoefenen van gezag en controle (planning, monitoring en handhaving) over het beheer van data assets. Dit is de basis en stuurt alle andere data managementfuncties aan. Het doel van data governance is dan ook om ervoor te zorgen dat data, in overeenstemming met beleid en best practices, op juiste manier worden beheerd. Het focust op hoe besluiten over data gemaakt worden en hoe mensen en processen verwacht worden te gedragen met betrekking tot data. De focus van een data governance programma is afhankelijk de behoeftes van de organisatie, maar zal meestal wel betrekking hebben op de volgende onderwerpen:

- **Strategy (strategie):** Het definiëren van, communiceren met en aansturen van de uitvoering van Data Strategy en Data Governance
- **Policy (beleid):** Het opstellen en handhaven van beleid met betrekking tot data en Metadata management, access, usage, security en quality
- **Standards and quality (standaarden en kwaliteit):** Het instellen en handhaven van normen voor data kwaliteit en data architectuur.
- **Oversight (toezicht):** Het bieden van hands-on observaties, audit en correcties op belangrijke gebieden van kwaliteit, beleid en data management (data stewardship).
- **Compliance (naleving):** Ervoor zorgen dat de organisatie kan voldoen aan de data gerelateerde wettelijke nalevingsvereisten
- **Issue management (probleem management):** Het identificeren, definiëren, escaleren en oplossen van problemen met betrekking tot data security, data access, data quality, naleving van regelgeving, data ownership, beleid, normen, terminologie of data governance procedures
- **Data management projects (data management projecten):** het steunen van pogingen op data management te bevorderen
- **Data asset valuation (waardering van data assets):** Standaarden en processen instellen om de business waarde van data te definiëren

Het is daarom belangrijk dat middels data governance policies en procedures ontwikkeld worden. Daarnaast is het ook belangrijk, dat het zich bezighoudt met organisatieveranderingsmanagement zodat dit ervoor zorgt dat de organisatie actief geïnformeerd wordt over de voordelen van verbeterde data governance en zo ook het gedrag heeft om data succesvol te beheren als een asset.

Om data dus succesvol te beheren als een bedrijfsasset is, is het ook belangrijk dat het bedrijf zelf de waarde van en datamanagement leert. Zelfs met de beste datastrategie, data governance en datamanagement plannen zal het niet slagen tenzij de organisatie de nodige veranderingen accepteert en beheert. De volgende principes kunnen helpen bij het leggen van een sterke basis voor data governance.

- Succesvolle data governance start met visionair en betrokken leiderschap. Datamanagement activiteiten worden bestuurd door een data strategie dat weer wordt aangestuurd door een enterprise business strategie.
- Data Governance is een business programma, daarom moet het zowel IT besluiten gerelateerd aan data beheren als de business interactie met data.
- In alle kennisgebieden van datamanagement is data governance een gedeelde verantwoordelijkheid tussen business stewards en technische datamanagement professionals.
- Data governance vindt door de hele organisatie plaats, op elk niveau.
- Omdat activiteiten op het gebied van data governance samenwerking tussen functionele gebieden vereisen, moet het data governance programma moet er een framework vastgesteld worden dat verantwoordelijkheden en interacties definieert.

Nu dat er een beeld is van wat er binnen de organisatie nodig om een goede data governance te hebben is de vraag nog, waarom. De twee hoofdredenen: risico's verminderen en processen te verbeteren.

Data governance kent ook verschillende activiteiten, hieronder worden ze uitgelegd:

Definiëren van data governance voor de organisatie: Het is hierbij de bedoeling dat het duidelijk wordt waar data governance over gaat, dus wát wordt er beheerd, wíe wordt beheerd en wie beheert het. Dit moet allemaal duidelijk zijn om het succesvol te maken. Dit kan gedaan worden door de gehele organisatie in plaats van naar bepaalde gebieden. Definiëren van de scope van data governance houdt meestal in dat het bedrijf zelf ook gedefinieerd wordt.

Uitvoeren van een beoordeling van gereedheid: Hierbij wordt de huidige status van de informatiemanagement capaciteiten beschreven. Volwassenheid en effectiviteit zijn cruciaal voor het plannen van een data governance programma. Dit houdt meestal in: begrijpen wat de organisatie doet met data; meet de huidige datamanagementmogelijkheden en capaciteiten. Daarnaast is het ook belangrijk om de mate van verandering binnen een organisatie te meten om er zo achter te komen waar weerstand kan ontstaan. Een beoordeling over samenwerking tussen management en datagebruikers is ook essentieel binnen deze activiteit. Er moet samengewerkt kunnen worden omdat dit meerdere domeinen betreft binnen een organisatie. Als er binnen een organisatie niet goed samen gewerkt kan worden kan dit een obstakel zijn voor stewardship. Als laatste zal er ook een Business alignment assessment gedaan moeten worden. Hierbij wordt er gekeken naar hoe in lijn de organisatie is met hen business strategie, met betrekking tot data gebruik.

Uitvoeren van discovery en business alignment: Het data governance -programma moet bijdragen aan de organisatie door specifieke voordelen te identificeren en te realiseren. Discovery activiteiten zullen de effectiviteit van het bestaande beleid en richtlijnen identificeren en beoordelen. Discovery kan ook kansen voor data governance identificeren om zo de bruikbaarheid en inhoud van data te verbeteren. Daarnaast is ook datakwaliteit in onderdeel van discovery. Beoordelen van datakwaliteit zorgt ervoor dat er meer inzicht komt in de bestaande problemen en obstakels, maar ook zal het inzicht geven op wat de impact is van een slecht datakwaliteit en de risico dat men dan loopt.

Ontwikkelen data governance strategie: Een data governance-strategie definieert de scope en aanpak van data governance. De data governance moet uitgebreid gedefinieerd worden en in relatie gebracht worden tot de algemene bedrijfsstrategie, de datamanagementstrategie en de IT-strategie.

Ontwikkelen van goals, principes en policies: De ontwikkeling van doelen, principes en beleid afgeleid van de Data Governance-strategie zal leidend zijn voor de organisatie in de gewenste toekomstige staat. Deze worden opgesteld door datamanagement professionals of door business beleidsmedewerkers.

Aanmoedigen van veranderingsmanagement: Voor veel organisaties verschilt de 'werkwijze' binnen data governance met de huidige situatie in hen organisatie. Om data governance op te nemen in de

organisatie zullen werkwijze/het gedrag van mensen moeten veranderen. Verandering is niet altijd even makkelijk en daarom dat veranderingsmanagement aangemoedigd moet worden.

Aan probleem management doen: Issue management is het proces voor het identificeren, kwantificeren, prioriteren en oplossen van data governance-gerelateerd problemen, waaronder vragen met betrekking tot rechten en procedures, problemen die komen door veranderingsmanagementprocessen, problemen omtrent het voldoen aan compliance-eisen, privacy- en vertrouwelijkheidsproblemen en problemen gerelateerd aan policies, standaarden, architectuur en procedures.

Beoordelen van Wettelijke nalevingsvereisten: Elke organisatie wordt beïnvloed door overheids- en sectorvoorschriften, inclusief voorschriften die bepalen hoe data en informatie moet worden beheerd. Een deel van de van data governance heeft te maken met het bewaken en waarborgen en naleving van de regelgeving. Naleving van regelgeving is vaak de eerste reden voor het implementeren van data governance.

Implementeren van data governance: Data governance kan niet van de ene op de andere dag worden geïmplementeerd. Het vereist een planning - niet alleen om verantwoording af te leggen over verandering, maar ook simpelweg omdat het veel complexe activiteiten omvat die gecoördineerd moeten worden. Implementatie-roadmap die tijdschema's voor, en relaties tussen verschillende activiteiten schetst is hierbij dan handig.

Ontwikkel een business glossary: Een glossary (woordenlijst) is nodig omdat mensen woorden anders gebruiken. Het is vooral belangrijk om duidelijkheid te hebben over definities van de data. Een glossary helpt hier dan bij.

Werk samen met architectuur groepen: Het business datamodel moet gezamenlijk worden ontwikkeld en onderhouden door data-architecten en Data Stewards. Afhankelijk van de organisatie kan dit werk worden gecoördineerd door de Enterprise Data Architect of door de steward.

Sponsor Data Asset Valuation; De waarde van data moet gemeten worden. Er zou een standaard moeten komen waaraan data moet voldoen.

Integreer data governance: Het is handig om een plan te hebben met betrekking tot de uitvoering van data governance. Hierin zou een lijst moeten opgenomen worden waarin alle activiteiten zijn weergegeven die nodig zijn voor het implementeren en uitvoeren van data governance activiteiten. Het schetst activiteiten, timing en technieken die nodig zijn om succesvol te zijn. Daarnaast is het ook handig om een data governance community of interest op te zetten. Vooral voor het eerste jaar is dit handig. Mensen kunnen zo beter weten wat het precies inhoudt en van elkaar leren.

Om uiteindelijk data governance te implementeren raad DAMA DMBOK aan om een onlineprogramma te gebruiken. Dit zorgt ervoor dat alle document op een centrale plek staan. Daarnaast brengen websites nog handige voordelen met zich mee. Zo kunnen bij websites documentation libraries aanwezig zijn en kunnen dingen snel opgezocht worden.

Data architecture: Een architectuur beschrijft het systeem binnen een organisatie. Het weergeeft alle componenten binnen het systeem, de relaties tot elkaar en tot de omgeving en de principes die het ontwerp en de verandering bepaald. Het is belangrijk om een goede architectuur te hebben. Dit helpt namelijk bij het begrijpen van de huidige systemen, naleving van regelgeving en maakt het effectiever. Data architectuur wordt ook wel gezien als de brug tussen de business strategie en de technologische uitvoering. Het stemt business en IT op elkaar af. Zo worden bijvoorbeeld eisen vanuit de business vertaald in data en systeem requirements om zo ervoor te zorgen dat bedrijfsprocessen uiteindelijk de juiste data binnen krijgen. Daarnaast helpt het ook bij het beheren van complexe gegevens- en informatielevering door de gehele organisatie heen.

Een data architectuur kent eigenlijk vier verschillende domeinen: Enterprise business Architecture, Enterprise Data Architect, Enterprise Applications Architecture en Enterprise Technology Architecture.

Het is belangrijk dat hierbij domeinen (en dus de architecten) gezamenlijk kijken naar ontwikkelingen, richtingen en vereisten. Elk domein heeft namelijk op elkaar. In onderstaande tabel is er per domein gekeken naar het doel, het element waaruit het bestaat, afhankelijkheden en de rollen binnen zo'n domein.

Tabel 9: Enterprise architectuur domeinen (DMBOK2, 2017)

<i>Domein</i>	<i>Enterprise business Architecture</i>	<i>Enterprise Data Architect</i>	<i>Enterprise Applications Architecture</i>	<i>Enterprise Technology Architecture</i>
<i>Doel</i>	Om te identificeren hoe een Organisatie waarde creëert voor klanten en andere belanghebbenden	Om te beschrijven hoe gegevens moeten worden georganiseerd en beheerd	Om de structuur en functionaliteit van applicatie in een onderneming te beschrijven	Om te omschrijven wat (fysieke) technologie nodig heeft om systemen op goede wijze te laten functioneren en zo waarde levert.
<i>Onderdelen</i>	Bedrijfsmodellen, processen, mogelijkheden, diensten, evenementen, strategieën, vocabulaire	Data modellen, definitie van gegevens, data flows, data Mapping Specification gestructureerde data API's	Bedrijfssystemen, softwarepakketten, databases	Technologische platformen, netwerken, beveiliging, integratietools
<i>Afhankelijkheden</i>	Stelt requirements voor andere domeinen vast	Beheert data gemaakt en vereist door de business architecture.	Handelt op specifieke data volgens de business requirements.	Hosten en uitvoeren van Application Architecture
<i>rollen</i>	Business architecten, business analisten, business data stewards	Data architecten, data steward en modelleurs	Applicatie architecten	Infrastructuur architecten

Dit kent twee activiteiten:

Vaststellen van data architectuur: Een data architectuur zou een integraal onderdeel moeten zijn van de Enterprise architectuur. Hierbij wordt er een framework gekozen dat van toepassing is op het bedrijfstype (gebruik bijvoorbeeld een overheid framework voor een overheidsinstantie). De opvattingen in het framework moeten bruikbaar zijn in de communicatie naar de verschillende belanghebbenden. Dit is vooral belangrijk voor Data Architecture-initiatieven, omdat ze betrekking hebben op bedrijven en systemen. Data-architectuur heeft een nauwe samenhang met bedrijfsarchitectuur

Integreren met Enterprise Architecture: Om Enterprise Data Architecture specificaties te ontwikkelen moet het eerst van een ontwerpniveau naar een meer gedetailleerd niveau dat in relatie is met de andere architectuur domeinen. Dit gebeurt meestal door middel van gefinancierde projecten. Gefinancierd projecten sturen over het algemeen architectuurprioriteiten aan. Toch moeten ook bedrijfswijde data architecture kwesties aangepakt worden. Data architecture kan zelf de scope van een project beïnvloeden. Het is daarom handig om Enterprise Data Architecture-kwestie te integreren met project portfolio management. Hierdoor kan een roadmap geïntegreerd worden en draagt het bij

aan een beter projectresultaat. Andersom moet ook de Enterprise data architecten betrokken worden bij de ontwikkeling integratie van Enterprise application.

7.6 Wat zegt De Cloud Data Mangement Capabilities Framework over het onderhouden/beheren van data(modellen)

Aanpak

Een ander framework waarnaar gekeken is, is het Cloud Data Management Capabilities Framework. Het is de bedoeling dat dit framework wordt vergeleken met DAMA DMBOK. Daarom dat ook hier voornamelijk de focus ligt op governance en architectuur. De reden dat hier ook naar is gekeken is omdat Rijkswaterstaat onder andere met TOPdesk werkt en ook de database gebruikt die achter TOPdesk zit. Dit is een Oracle-database, daarnaast maken ze ook gebruik van een Postgres en een MS SQL server (Rijkswaterstaat, 2018). Dit zijn Cloud based databases, daarom dat dit misschien ook een handig framework voor Rijkswaterstaat kan zijn.

Resultaten/uitleg

Het Cloud Data Mangement Capabilities (CMDC) Framework definieert de best practice mogelijkheden die nodig zijn om data te managen en te controleren in een Cloud omgeving. Het doel van dit framework is om onder andere vertrouwen en betrouwbaarheid op te bouwen in Cloud technologie. CMDC is een best practice assessment en framework voor het manage en controleren van data in single, multiple en hybride Cloud omgevingen. Het wordt gebruikt om de capaciteiten van een organisatie, die nodig zijn voor het ondersteunen van een gecontroleerde integratie en migratie naar de Cloud omgeving te beoordelen. Het focust zich op, en bereidt zich uit op de mogelijkheden die cruciaal zijn voor het controleren van belangrijke en gevoelige data, en benadrukt kenmerken van hedendaags Cloud platforms die kansen bieden voor standaardisatie en automatisering van databeheer en -controle. (EDM Council Inc, 2021)

Om data in een Cloud omgeving te beheren zijn er een paar datamanagement capabilities:

- Een organisatie moet duidelijke de verantwoording, beheer en governance vaststellen voor data die zijn gemigreerd naar of gemaakt in Cloudomgevingen
- Het is een vereiste dat men te allen tijde weet wat voor data er in de Cloud omgeving zit en hoe gevoelig deze data is. Dit bijhouden is essentieel om automatisch beheer van datagebruik en data toegang.
- Datamanagement beheer moet vastgesteld worden aan de hand van de levenscyclus van de data
- Data assets moeten geschikt zijn voor de beoogde doelen en zich houden aan de eisen van bewaring en archivering.
- Net als bij on-premises data-assets, is het ontwerp van de data-architectuur en de configuratie van ondersteunende technologieën belangrijk om ervoor te zorgen dat de bedrijfsdoelstellingen worden gehaald. (EDM Council Inc, 2021)

CDMC legt de vereisten voor deze capabilities vast in zes gebieden. Deze zes componenten van het frameworks, bevatten in totaal 14 capabilities en 37 sub- capabilities. Dit framework kent de volgende zes gebieden: Governance & Accountability, Cataloging & Classification, Accessibiliy & Usage, Pretection & Privacy, Data lifecycle en Data & Technical Architecture. Om binnen de scope te blijven wordt nu alleen gekeken naar Governance & Accountability en Data & Technical Architecture.

Governance & Accountability: Dit gebied vormt de basis voor succesvol datamanagement in Cloudomgevingen. De Cloudomgeving introduceert uitdagingen en kansen voor schaalvergroting, standaardisatie, automatisering en het gedeelde verantwoordelijkheidsmodel. Daarom is het belangrijk om een effectief data governance programma toe te passen op data die zich in een Cloud omgeving bevindt. Alle belanghebbenden moeten een duidelijk begrip hebben van gegevenscontroles en verantwoordelijkheid voor elke rol. De aanpak is vergelijkbaar met hoe data governance, -controls en -accountability zijn toegepast op gebruikelijk datamanagement in een organisatie.

Hierbij is de scope als volgt:

- Definieer business cases voor het managen van data in de Cloud.
- Zorg ervoor dat rollen en verantwoordelijkheden van data-eigenaren zich ook uitbreiden naar data in de Cloud.
- Zorg ervoor dat databronnen beheerd wordt aan de hand van goede bronnen en distributeurs.
- Automatisering van de Cloud zorgt voor kansen met betrekking tot data consumption.
- Begrijp requirements voor het beheren van datasoevereiniteit en de risico's van cross-border data movement.
- Implementeer controles voor datasoevereiniteit en cross-border data movement (EDM Council Inc, 2021)

(Datasoevereiniteit heeft te maken met de data van de burgers in het land. De overheid probeert te voorkomen dat deze gegevens in de verkeerde handen vallen, dit doet zij door maatregelen die ervoor moeten zorgen dat bedrijven persoonsgegevens niet in het buitenland vrij laten.)

Data & Technical Architecture: Data & Technical Architecture gaat over de architectonische problemen die uniek zijn voor de Cloud en hoe dit het managen van data in de Cloud omgeving beïnvloed. De meeste Cloud serviceproviders bieden verschillende opties aan voor hoe business oplossingen middels verschillende Clouds service ontworpen en geïmplementeerd kunnen worden. Het ontwikkelen en toepassen van bepaalde architectuurpatronen en richtlijnen kunnen zorgen voor een basis met betrekking tot data management in de Cloud. (EDM Council Inc, 2021)

Dit gebied bevat een aantal capabilities om ervoor te zorgen datastromen in, uit of tussen Cloudomgevingen begrepen worden en dat architectuur richtlijnen helpen bij het ontwerpen ontwerp van Cloud computing-oplossingen.

Het gebied focust op de volgende dingen:

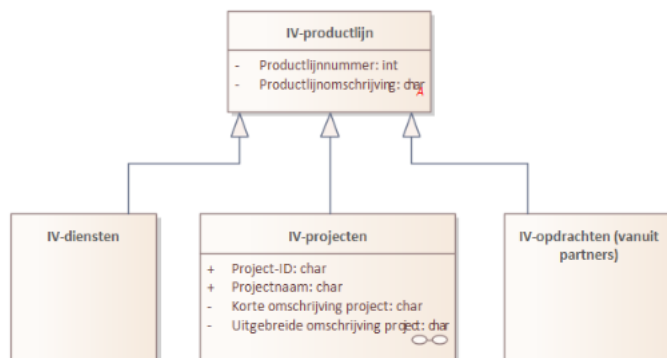
- Het vaststellen en toepassen van principes voor beschikbaarheid van data.
- Ondersteunen van business requirements voor backups en herstellen van data
- Optimaliseren van het gebruik van de Cloud en de bijbehorende kosten voor de service.
- Ondersteunen van data portability en de mogelijkheid om data tussen verschillende Cloud omgevingen te verplaatsen
- Automatiseren van identificeren van dataprocessen en datastromen in de Cloud omgeving zelf en tussen verschillende Cloud omgevingen.
- Het identificeren, volgen en managen van verandering op gebied van oorsprong van de data. Dit maakt het mogelijk om de oorsprong van data altijd te kunnen achterhalen en uit te leggen.
- Tools die op zinnige wijze rapporteren en visualiseren, zowel vanuit een business perspectief als een technisch perspectief (EDM Council Inc, 2021)

8. BEROEPSPRODUCT

Tijdens mijn stage zijn er twee producten tot stand gekomen. Een datamodel met betrekking tot de IT-projecten en een advies met betrekking tot hoe dit model het best te onderhouden is.

8.1 De datamodellen

Verantwoording – datamodel bedrijfsvoering IV: De CIV onderkent vijf productlijnen/productiestromen. Op basis van “opdrachtgeverschap” en “waar komt de financiering vandaan” zijn deze productlijnen tot stand gekomen. Elk project type is te relateren aan één productlijn. Er wordt niet verder ingegaan op productlijn, maar de productlijn bestaat uit behalve projecten ook een IV-diensten en IV-opdracht. Deze zijn in dit model niet verder uitgewerkt omdat dat buiten de scope valt, maar wel gemodelleerd om zo een volledig beeld te schetsen en zo een mogelijkheid te creëren om zo een koppeling te maken naar andere datamodellen.



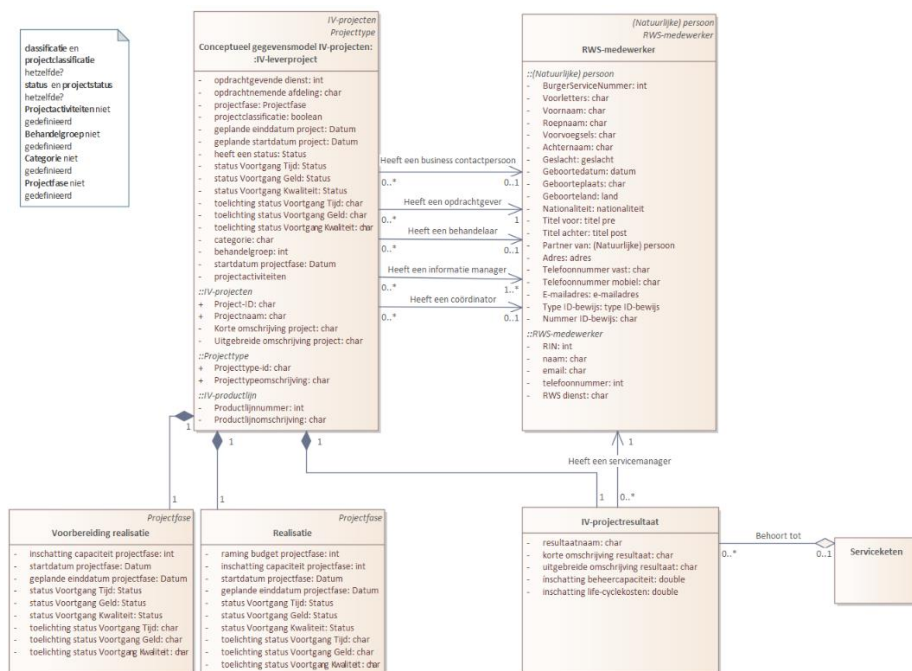
Figuur 14: Het datamodel bedrijfsvoering IV

Verantwoording – datamodel IV-projecten: Het IV-projecten datamodel ziet er zoals bovenstaande afbeelding uit. Er wordt onderscheidt gemaakt tussen verschillende type projecten. Er is daarom een parent klasse: IV-project aangemaakt die algemene informatie van de projecten bevat. Deze is vervolgens ‘opgesplitst’ in de drie soorten projecten: IV-ontwikkelpject, IV in GWW-project en IV-leverproject. In dit model is per projecttype de bijhorende attributen weergegeven, maar deze worden in andere modellen verder uitgewerkt. De reden dat ik ervoor heb gekozen om het zo aan te pakken is omdat het anders een te groot model wordt en daardoor niet meer overzichtelijk wordt. Door nu niet elk type project in detail te modelleren blijft het overzichtelijk. Onder elke type project klasse hangt een ander model die hier verder op in gaat. Daarnaast is het zo ook makkelijker om nieuwe klassen aan het model toe te voegen. Stel er komt een nieuw type project bij dan kan dat aan dit model toegevoegd worden, zonder dat dit andere klassen beïnvloed. Daarna kan het nieuwe type project verder uitgewerkt worden in een ‘eigen’ model, maar hangt het wel vast aan dit model.

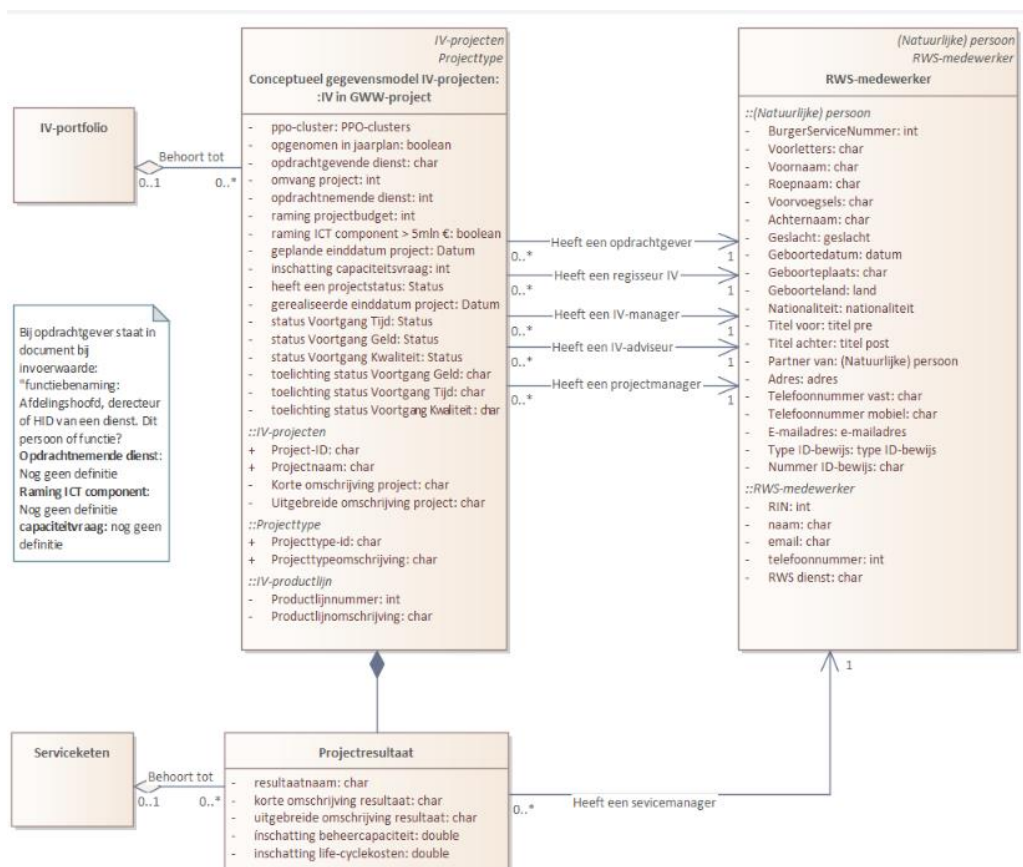
EA heeft de mogelijkheid om klassen als het ware ergens anders ook te gebruiken. Dit zorgt ervoor dat de attributen, operaties, notities van een klasse in diagram 1 hetzelfde is als in diagram 2. Ook wanneer dit veranderd wordt, zullen de klassen hetzelfde blijven. Hierdoor hoeft er maar op één plek een klasse aangepast te worden.



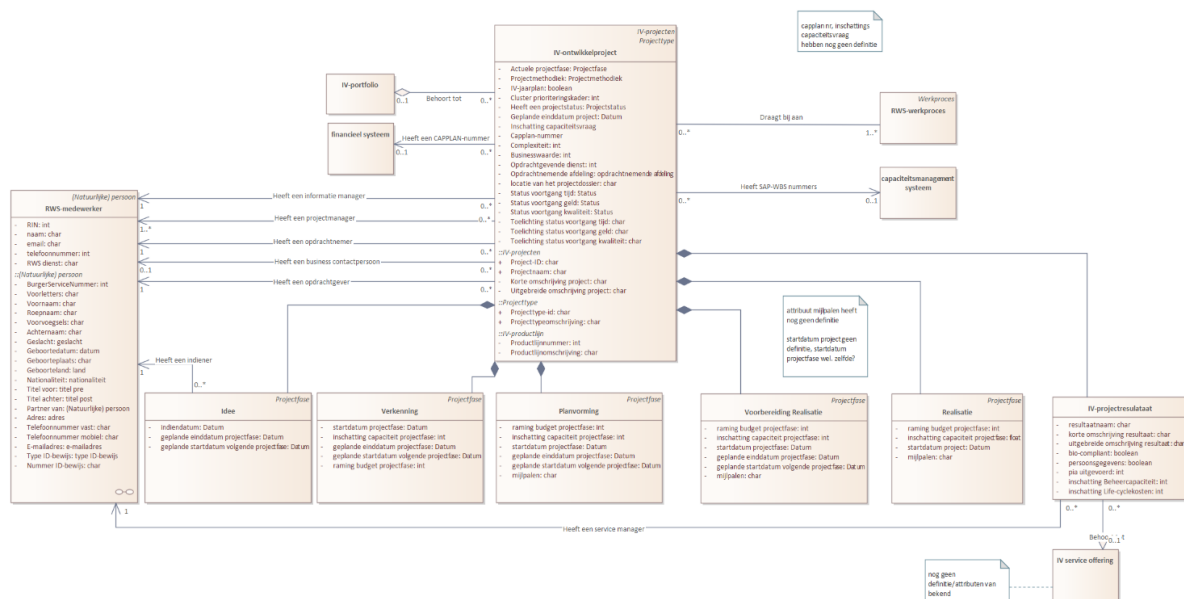
Pagina | 37



Figuur 16: Het datamodel IV-Leverproject



Figuur 17: Het datamodel IV in GWW-project



Figuur 18: Het datamodel IV-ontwikkelproject

8.2 Het advies

De vraag was hoe de data(modellen) het beste onderhouden/beheerd kunnen worden. Het antwoord erop was echter niet zo makkelijk. Er was weinig literatuur te vinden op gebied van onderhoud datamodellen. Om uiteindelijk toch tot een goed advies te komen heb ik daarom naar twee dingen gekeken:

1. Tijdens het modelleren van de data: hoe kan dit vervolgens het makkelijkst onderhouden worden. Hierbij is er voornamelijk gekeken naar/vanuit de datamodellen.
2. Het probleem achter de datamodellen was als volgt: definities zijn onduidelijk → ingevoerde data heeft daardoor mogelijk niet dezelfde betekenis → gegevens op het dashboards zijn incorrect. Datamodellen het definitie probleem oplossing, echter is het nog wel de vraag of de ingevoerde data voldoet aan de beschrijving (de definitie) van een klasse. Is dit niet het geval dan heeft dit nog steeds gevolgen voor de dashboards. Om deze reden is er ook gekeken naar data management en met name data governance. Data governance zou er namelijk voor kunnen zorgen dat dit probleem opgelost wordt. Hiervoor is gekeken naar twee frameworks.

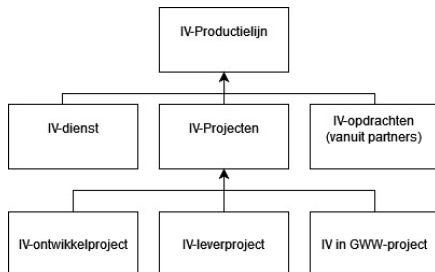
Om uiteindelijk tot een compleet mogelijk advies te komen is er dus gekeken naar de datamodellen en naar twee datamanagement frameworks. Er is gekeken naar het DAMA DMBOK en het CDMC™ (Cloud Data Management Capabilities) framework. Waarom er is gekozen voor deze twee frameworks en wat ze inhouden is in het volgende hoofdstuk te lezen.

8.2.1 Datamodellen ontwerp

8.2.1.1 Het gekozen ontwerp

Bij het ontwerpen van de datamodellen is er voor gekozen om het in verschillende niveaus op te delen. Hierdoor is er eerst naar de bovenste laag gekeken dus de bedrijfsvoering. Hierin is te zien waar de IV-productielijn uit bestaat. In een nieuw diagram is er gekeken naar IV-projecten zelf. Hierbij is er gekeken naar wat er in het algemeen bij een IV-project komt kijken. Denk aan bijvoorbeeld budget, ongeacht welk type project hoort er altijd wel een budget bij. Hier zijn ook de 3 verschillende type projecten weergegeven, maar deze zijn verder niet uitgewerkt. Dit wordt gedaan op het volgende niveau. Hierbij is elk project afzonderlijk uitgewerkt. Klassen komen hierdoor dus wel in meerdere diagrammen voor, maar omdat er een kopie van een klasse gemaakt wordt is het dus niet het geval dat je dezelfde klasse op verschillende plekken moet bijwerken zodat deze ook in de andere diagrammen hetzelfde zijn. Bij het bewerken van een klasse, zullen alle kopieën ook automatisch

bijgewerkt worden. Dus je past bijvoorbeeld de klasse IV-Projecten aan in het diagram IV-productielijn. Deze klasse staat ook vermeld in de diagram IV-projecten, maar omdat deze klasse in de diagram IV-project een kopie is hoeft deze dus niet apart aangepast te worden. Hieronder is een schematische weergave van hoe de datamodellen met elkaar zijn verbonden.



Voordelen van deze manier:

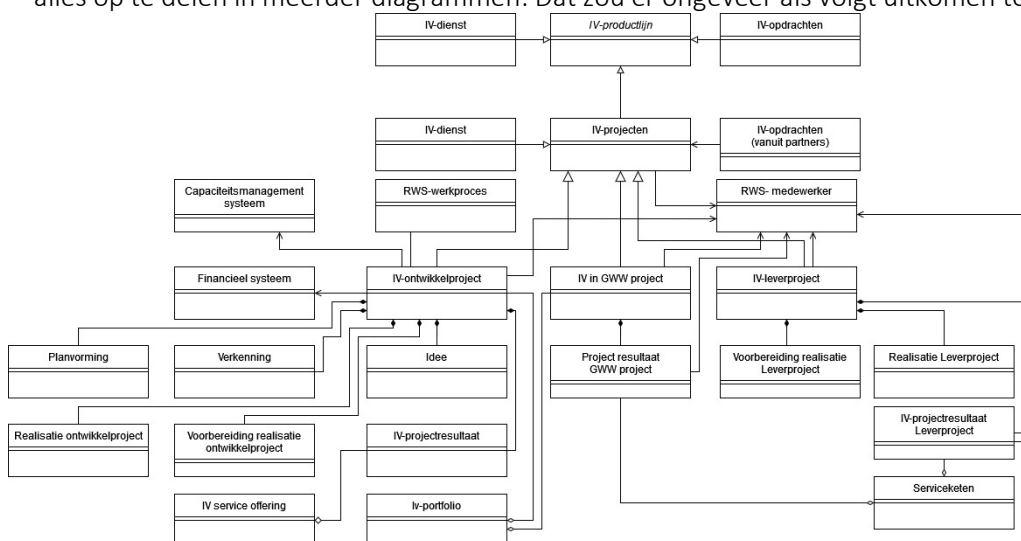
- Wanneer er een nieuw type project bijkomt dan kan dit in een apart diagram gemaakt worden waarin de project fases uitgewerkt worden. De hoofdklasse moet dan alleen toegevoegd worden aan het datamodel: IV-projecten. Het bijwerken van het model is hierdoor erg eenvoudig.
- Bij het implementeren van bijvoorbeeld een nieuw type project zal dit geen invloed hebben op de rest.
- Het is overzichtelijk, er staan niet 30 klassen in één model.

Nadelen van deze manier:

- Uiteindelijk kunnen er veel verschillende diagrammen ontstaan
- De mappenstructuur van hoe de diagrammen zijn opgeslagen is cruciaal om een logisch en hiërarchisch beeld te krijgen.
- Vanuit de bovenste twee lagen is er niet in één oogopslag te zien waar een project uit bestaat. Men zal hier eerst naar het juiste diagram moeten gaan

8.2.1.2 Het alternatief op het gekozen ontwerp

Echter was er ook een alternatief. Eventueel kon alles ook in één diagram komen te staan in plaats van alles op te delen in meerder diagrammen. Dat zou er ongeveer als volgt uitkomen te zien:



Voordelen van deze manier

- In één diagram te zien waar een project bijvoorbeeld uit bestaat en hoe het is verbonden met andere projecten, zonder andere diagrammen erbij te hoeven halen.
- Klassen komen ook maar één keer voor in het hele diagram. Hier is bijvoorbeeld zo te zien dat twee van de drie projecten een relatie hebben met serviceketen.

Nadelen van deze manier

- Veel Klassen waardoor het onoverzichtelijk wordt. Het duurt even voordat men alle entiteiten in zich heeft opgenomen.
- Heel veel relaties die vaak door elkaar heen lopen.
- Het toevoegen van een nieuw type project kan lastig worden omdat er ruimte gemaakt moet worden voor de klassen en de relaties.
- Het is even zoeken voordat je ziet welke klassen bij een project horen.

8.2.1.3 Conclusie/aanbeveling

Uiteindelijk is het gekozen ontwerp het meest praktische. Door dat het is opgedeeld in verschillende diagrammen is het kleiner en overzichtelijker. Het is ook makkelijker om nieuwe projecttypes/klassen aan bestaande projecttypes eraan toe te voegen omdat dit niet in een al volle diagram erbij gezet moet worden. Er kan namelijk gewoon een aparte diagram gemaakt worden voor een nieuw project of klassen aan een bestaand diagram toegevoegd worden. Als het aan een bestaand diagram toegevoegd moet worden dan gaat dit makkelijker in het huidige ontwerp dan wanneer alles in één diagram staat. Ook als er gekeken wordt naar UML best practices (Bellekens, 2018), dan komt het huidige ontwerp beter overeen met de punten die hierin genoemd zijn. De huidige manier heb ik ook besproken met Victor van der Kloet, en hij was het hier mee eens.

8.2.2 Data management framework

8.2.2.1 Requirements

8.2.2.2 DAMA DMBOK

Er is naar dit framework gekeken omdat het een van de bekendste data management frameworks is. Daarnaast kwam ook de vraag vanuit de organisatie om hiernaar te kijken en is ook bij de documentatie van de opzet van de conceptuele datamodellen aangegeven dat dit handig is om naar te kijken. Om deze redenen is er dus gekeken naar dit framework

Wat houdt het in:

DAMA DMBOK framework wordt gezien als een van de bekendste frameworks op gebied van datamanagement. Dit boek gaat over 11 kennisgebieden waarbij het draait om databeheer (SynTouch, 2020) (Bluefrog B.V., 2021) (Dama International et al., 2017). Het kent de volgende 11 kennisgebieden: Data governance, Data architecture, Data modeling & Design, Data Storage & Operations, Data security, Data Integration & Interoperability, Documents & Content, Reference & Master Data, Data Warehousing & Business Intelligence, Metadata, Data Quality. Voor een uitgebreidere uitleg op de inhoud van dit framework [zie hoofdstuk 7.1.5](#).

Voordelen van deze manier:

- Zeer theoretisch maar gaat ook in op de praktische kant
- Niet geheel onbekend voor de organisatie
- Als alle activiteiten uitgevoerd worden komen er veel documenten tot stand.
- Heel erg gericht op verandermanagement

- Voor iedereen beschikbaar
- Een zeer bekend framework voor datamanagement, bekender dan CMDC (Steenbeek, 2021)

Nadelen van deze manier:

- Het is mogelijk te uitgebreid. Er is nu naar twee kennisgebieden gekeken, maar dit framework kent er veel meer.

8.2.2.3 Cloud Data Management Capabilities Framework

Rijkswaterstaat kent drie soorten databasesystemen: Oracle-database, Postgres en een MS SQL server (Azure). De reden dat ze drie verschillende databasesystemen hebben is omdat elk systeem verschillende (integratie) eigenschappen kent. Een MS SQL server past het best bij een Microsoft omgeving. Andere systemen kunnen hier ook mee werken, maar dit gaat niet heel makkelijk. Om deze reden hebben ze dus 3 verschillende databasesystemen (Rijkswaterstaat, 2018). Zo'n systeem kent verschillende databases en binnen/tussen de databases kan data verplaats, gecontroleerd en beheerd worden. Het Clouw Data Mangement Capabilities (CMDC) Framework gaat hier ook voornamelijk op in en kijkt hoe dit goed gedaan kan worden tussen verschillende cloud omgevingen.

Wat houdt het in:

Het Clouw Data Mangement Capabilities (CMDC) Framework definieert de best practice mogelijkheden die nodig zijn om data te managen en te controleren in een Cloud omgeving. Het doel van dit framework is om onder andere vertrouwen en betrouwbaarheid op te bouwen in Cloud technologie. CMDC is een best practice assessment en framework voor het manage en controleren van data in single, multiple en hybride Cloud omgevingen. Het wordt gebruikt om de capaciteiten van een organisatie, die nodig zijn voor het ondersteunen van een gecontroleerde integratie en migratie naar de Cloud omgeving te beoordelen. Het focust zich op, en bereidt zich uit op de mogelijkheden die cruciaal zijn voor het controleren van belangrijke en gevoelige data, en benadrukt kenmerken van hedendaags Cloud platforms die kansen bieden voor standaardisatie en automatisering van databeheer en -controle. (EDM Council Inc, 2021).

Voor een uitgebreidere uitleg op de inhoud van dit framework [zie hoofdstuk 7.6](#).

Voordelen van deze manier:

- Heel erg cloud gericht, meer dan DAMA DMBOK
- Gaat vooral in op hoe data binnen en tussen cloudomgevingen gemanaged en gecontroleerd kan worden.
- Praktische focus

Nadelen van deze manier:

- Wordt verwacht dat de organisatie een sterke basis heeft op gebied van datamanagement
- Zeer onbekend voor de organisatie
- De kijk op implementatie van datamanagement ontbreekt
- Afgeleid van DCAM, DCAM is niet toegankelijk voor iedereen. Daarom dat ze dit hebben gepubliceerd. Hierin wordt voornamelijk de aanpak besproken voor dit framework.
- Om volledige toegang te krijgen tot dit framework is een organisatie lid worden van the EDM council, dit is redelijk prijzig.

8.2.2.4 Conclusie/aanbeveling

Er is dus gekeken naar twee verschillende frameworks; DAMA DMBOK en Cloud Data Management Capabilities Framework. Het advies hierbij is om vooral het DAMA DMBOK framework aan te houden.

Hier zijn meerdere redenen voor. Ten eerste wordt bij het Cloud Data Management Capabilities Framework al verwacht dat de organisatie een sterke basis heeft op gebied van datamanagement. Dit framework gaat daar namelijk nog veel dieper op in. De vraag is dan, is op het moment datamanagement binnen Rijkswaterstaat al zover, heeft het al zo'n sterke basis? Binnen Rijkswaterstaat zijn er een paar documenten aanwezig over data governance, maar de datastrategie is ook nog maar twee jaar oud. Mogelijk dat Rijkswaterstaat nog niet op het niveau zit om dit framework toe te passen. Daarnaast is er niet veel theorie omschreven in dit framework. Het heeft meer een praktische focus. Om die reden kan het ook moeilijker zijn voor beginners om dit te begrijpen.

DAMA DMBOK is echter weer heel erg theoretisch. Het verschaft veel informatie over datamanagementcomponenten, maar heeft ook een praktische kant. Heeft het ook over ervaringen en laat zien wat echt werkt. Een andere reden waarom DAMA DMBOK aan te raden is, is omdat dit in de organisatie al een beetje bekend is. In ieder geval meer bekend dan Cloud Data Management Capabilities Framework. Voordat ik kwam was er al aangegeven dat dit een goed framework was om te gebruiken, dit impliceert dat op z'n minst een paar mensen binnen de organisatie kennis hebben van DAMA DMBOK. Het is moeilijk om zaken binnen een grote organisatie als Rijkswaterstaat te veranderen. Er was namelijk al via de interviews naar voren gekomen dat zaken pas veranderen, als dit voor een probleem begint te zorgen. Het is daarom misschien ook niet verstandig om een geheel nieuw framework te introduceren, terwijl ze al een ander prima bruikbaar framework kennen.

Als er gekeken wordt naar de requirements en deze in een L-matrix uitgedrukt worden dan is het volgende te zien:

Tabel 10: L-matrix requirements DAMA DMBOK vs CDMC

<i>Requirement</i>	<i>DAMA DMBOK</i>	<i>Score</i>	<i>CDMC framework</i>	<i>Score</i>
De governance rollen worden ingevuld	Duidelijk, een van de eerste activiteiten binnen DAMA DMBOK – governance is het definiëren van wat er wordt beheerd en wie dit beheerd.	5	Matig, Het gaat meer over hoe rollen en verantwoordelijkheden van data-eigenaren zich ook uitbereid naar data in de Cloud	3
Er is een erkende eigenaar voor de data aangewezen	Duidelijk, dit framework gaat ook in op data ownership. Zo wordt er onderscheid gemaakt tussen verschillende data owners.	5	Duidelijk, een van de doelen is ook om rollen van data eigenaren te definiëren en dit uiteindelijk ook toe te passen in de cloud omgeving.	5
Totaal		10		8

Hierbij geldt: 5: voldoet in hoge mate, 3 Voldoet in zekere mate, 1, is enigszins van toepassing. (managementmodellensite, 2021)

Kortom, kijkend naar de organisatie, de requirements van de organisatie en hoever de organisatie is met data management is het aan te raden om vooral te kijken naar het DAMA DMBOK.

9. REFLECTIE

In dit hoofdstuk wordt er gereflecteerd op de stage, de competenties, de planning, aanpak, uitkomsten, risico's en begeleiding.

9.1 Competenties

9.1.1 A-Competenties

Hieronder is per A-competentie een beschrijving te vinden van wat dit inhoudt, vervolgens is er middels STARR methode gereflecteerd op de competentie

9.1.1.1 Onderzoek

Beschrijving: Onderzoek is een vertaling van de A-Competentie “A3 (Wetenschappelijke) Toepassing [onderzoek]” en “A7 Methodisch en Reflectief denken en handelen”, waarin we de nadruk willen leggen dat elk afstuderen een onderzoekstraject bevat waarin alle aspecten van het onderzoekproces terugkomen. Waar dit proces ingebracht wordt en de omvang ervan kan per opdracht en specialisatie anders zijn (Hogeschool Leiden, 2021).

Situatie: Rijkswaterstaat wilde hen conceptuele datamodellen in UML-format in EA implementeren en hadden hierbij ook de vraag hoe dit vervolgens het best onderhouden kan worden. De reden hiervoor was zoals in [hoofdstuk 3: Probleemstelling](#) beschreven staat, om zo onder andere de definities vast te leggen.

Taak: Om uiteindelijk een goed datamodel te maken en een goed advies te kunnen geven heb ik een onderzoek gedaan. Hierbij zijn de deelvragen bedoeld om uiteindelijk zowel een goed modellen als een goed advies op te kunnen leveren. Mijn taak bij het vore van de deelvragen was om ze zo smart mogelijk te maken.

Actie: Tijdens deze stage is er uitvoerig onderzoek gedaan naar onder andere Rijkswaterstaat zelf, UML en DAMA DMBOK. Dit is gedaan middels kwalitatief onderzoek waarbij ik ook verschillende methodes en technieken heb gebruikt. Zo heb ik MoSCoW-analyse gebruikt voor het analyseren van de requirements. Daarnaast moest ik ook vooral kijken naar interne en externe documenten voor zowel onderbouwing als het vergaren van kennis. Het doel het onderzoek was om zo de deelvragen uit te werken, dit zou leiden tot het beantwoorden van de hoofdvraag (zie hiervoor [hoofdstuk 5: Het onderzoek](#)) uit te werken. Ik heb hiervoor onder andere interviews gehouden, een workshop gedaan en boeken gelezen, om zo tot een antwoord te komen.

Resultaat: Uiteindelijk komt door het beantwoorden van de deelvragen en daarmee de hoofdvraag datamodellen met betrekking tot IT-projecten en een advies over hoe de datamodellen onderhouden kunnen worden tot stand. Er zijn datamodellen tot stand gekomen die beschrijvingen bevatten van relaties, klassen en attributen. Bij het modelleren is er rekening gehouden met eventuele aanpassingen die zich in de toekomst kunnen voordoen. Daarnaast is er ook een advies tot stand gekomen dat zowel kijkt naar het ontwerp van de datamodellen als naar twee verschillende frameworks. Voor het resultaat zie [hoofdstuk 8: beroepsproduct](#).

Reflectie: Ik ben van mening dat het onderzoek goed verlopen is. Er waren genoeg bronnen te vinden en uiteindelijk is iedere vraag ook beantwoord. Ik ben tevreden met het resultaat en heb hier veel van geleerd. De datamodellen staan namelijk in de ‘echte’ omgeving en er is een formeel advies gekomen. Dit alles is ook gelukt binnen het tijdsframe.

9.1.1.2 Leren Leren

Beschrijving:

‘Leren leren’ is een vertaling van de A-competenties “A1 brede professionalisering” en “A4 transfer en brede inzetbaarheid”, waarin we de nadruk willen leggen opdat elke student in staat moet zijn zichzelf te vernieuwen wanneer nodig. Dit is een essentiële competentie in het ICT-vakgebied dat met constante veranderingen te maken heeft. Als basis van ‘leren leren’ geldt het positieve zelfbeeld wat een student in staat stelt zelfredzaam te zijn en motiveert om nieuwe dingen op te pakken. Elke student moet sturing kunnen geven aan het zelfbeeld, moet zichzelf kunnen redden doordat hij de juiste tools kan toepassen en kan zichzelf motiveren om tot resultaten te komen (Hogeschool Leiden, 2021).

Situatie: Voor mijn stage had ik nog niet veel kennis van en ervaring met EA, houden van interviews, transcriberen van de interviews en werken met de Kanban methode. Gedurende mijn stage kwam ik met al deze dingen in aanraking.

Taak: Het doel was om meer kennis en uiteindelijk ervaring te hebben in de verschillende onderwerpen zodat ik dit uiteindelijk zelf ook kon doen.

Actie: Tijdens het project heb ik onderzoek gedaan naar de verschillende project methodes, uiteindelijk een trello bord opgezet, interviews gehouden, gekeken naar de verschillende manieren van transcriberen en de interviews getranscribeerd. Ook heb ik een workshop over EA gedaan. Voor deze stage had ik geen kennis van EA, maar dit is inmiddels verbeterd door de workshop en eigen onderzoek naar hoe EA werkt.

Resultaat: Doordat ik datamodellen heb opgesteld in EA, weet ik nu ook beter hoe het werkt. En doordat ik advies heb gegeven op datamanagement frameworks, kan ik nu ook met zekerheid zeggen dat mijn kennis hierin vergroot is. Daarnaast heb ik mijn interviews gehouden en deze getranscribeerd. Ook heb ik weer meer ervaring met verschillende methodes waar ik eerst nog nooit mee heb gewerkt zoals SERVQUAL-methode.

Reflectie: Doordat ik mij heb ingelezen in (nieuwe) onderwerpen (CDMC), hier advies op heb gegeven en ook oude kennis naar boven heb gehaald (UML), doordat ik weer data moest modelleren, kan ik zeggen dat ik veel heb geleerd. Daarnaast heb ik ook een nieuw programma geleerd, EA. Voor de stage had ik nog geen ervaring met de tooling, maar ik kan nu met zekerheid zeggen dat ik nu met EA overweg kan dan 5 maanden geleden.

9.1.1.3 Professioneel werken

Beschrijving: Professioneel werken is een vertaling van de A-competenties "A6 probleemgericht werken", "A7 methodisch en reflectief denken en handelen" en "A8 Sociaal-communicatieve bekwaamheid" waarin de aanpak van het afstuderen centraal staat. De aanpak is een rode lijn door het hele afstuderen, wat start bij het probleem en eindigt bij de reflectie achteraf (Hogeschool Leiden, 2021).

Situatie: Tijdens mijn stage heb ik besloten om een projectmethode toe te passen. Hiervoor heb ik dan ook een methode uitgezocht. Lees hierover meer in [hoofdstuk 6: Methodologie](#). Daarnaast werd er ook verwacht dat je een beetje meedraait in het werkproces, zo is er een wekelijks stand-up waarbij de afdeling verteld waar hij/zij mee bezig is. Ondanks dat ik altijd aan de afstudeeropdracht heb gewerkt, waren er ook workshops/evenementen/bedrijfsuitjes om rijkswaterstaat beter te leren kennen waar ik aan heb deelgenomen.

Taak: Het doel hiervan is om enerzijds een goed product op te leveren, anderzijds ook mee te draaien in het werkproces. Hiervoor moest ik zelf informatie verzamelen en stakeholders benaderen. Daarnaast werd er dus ook verwacht dat je een beetje meedraaide in het werkproces, hierbij was het natuurlijk ook de bedoeling dat er aan afspraken gehouden werd.

Actie: Ik heb gekeken naar mogelijke [projectmethodes](#) en ben nagegaan welke het beste bij mij paste. Verder voor het project heb ik ook gesprekken gehad met andere medewerkers. Dit waren zowel gewone kennismakingsgesprekken als [interviews](#). Ook heb ik deelgenomen aan de wekelijkse stand-up en aan stakeholder de modellen laten zien voor feedback. Deze heb ik ook vervolgens verwerkt.

Resultaat: Uiteindelijk ben ik gewoon een ‘medewerker’ geweest die individueel aan een project heeft gewerkt. Hiervoor heb ik contact opgenomen met belanghebbende en algemene workshops en evenementen vanuit Rijkswaterstaat bijgewoond.

Reflectie: De ervaring om 40 uur per week voor een bedrijf te werken was niet nieuw, maar wel de omgeving. Dit was weer een heel ander soort bedrijf dan mijn vorige stagebedrijf. Het had een andere dynamiek. Daarnaast heb ik nu ook gewerkt met een duidelijke planning.

9.1.1.4 Innovatie

Beschrijving: Innovatie, is boven de theorie staan, buiten de gestandaardiseerde reglementen en aanpak kunnen denken. Buiten de kaders dus. Mag inhoudelijk of op proces. Moet leiden tot iets nieuws, zoals een maker het betaamt. Innovatie is een vertaling van de A-competenties: “A4 Transfer en brede inzetbaarheid”, “A5 Creativiteit en complexiteit in handelen” en “A6 Probleemgericht werken” (Hogeschool Leiden, 2021).

Situatie: Zoals al in de [probleemstelling](#) naar voren kwam, is het concept datamodellen nog relatief nieuw binnen Rijkswaterstaat. Dit werd ook tijdens de interviews bevestigd. Rijkswaterstaat had amper tot geen eigen conceptuele datamodellen. De reden hiervoor is terug te lezen in [hoofdstuk 4: De context](#).

Taak: Er werd verwacht dat ik nieuwe datamodellen ging maken.

Actie: Om deze datamodellen te kunnen is er eerst gekeken naar hoe EA werkt. Vervolgens is er ook gekeken naar hoe dit het best in UML-format kan komen te staan. Uiteindelijk heeft dit ertoe geleid dat er nieuwe datamodellen tot stand zijn gekomen. Deze staan in UML-format en zijn geïmplementeerd in EA.

Resultaat: Uiteindelijk zijn er dus nieuwe datamodellen mb.t. IT-projecten in EA gekomen. Deze staan ook al in de echte omgeving (niet meer de omgeving waar de zaken eerst in komen), voor de datamodellen zie [hoofdstuk 8: Beroepsproduct](#). Deze modellen zijn nog niet eerder gemaakt en daarmee dus erg nieuw en innovatief.

Reflectie: Terug reflecterend op de stage ben ik van mening dat dit zowel voor mij als voor het bedrijf heeft geresulteerd in innovatie. Voor het bedrijf was dit een heel nieuw product, deze datamodellen waren er immers nog niet. Voor mij heb ik nieuwe dingen geleerd (Cloud Data Management Capabilities Framework, EA) die mij zullen helpen in de toekomst. Al met al kan ik zeggen dat deze stage veel nieuwe dingen heeft opgeleverd.

9.1.2 B-Competenties

Tijdens mijn stage heb ik gekozen om de volgende drie B-competenties aan te tonen:

- Organisatieprocessen analyseren – opdracht in context
- Organisatieprocessen adviseren – adviseren dataoplossing
- Infrastructuur analyseren – analyseren data architectuur

Ik heb voor deze 3 competenties gekozen, omdat deze in mijn optiek gedurende mijn stage aangetoond konden worden. Hieronder wordt per competentie uitgelegd uit welke onderdelen deze bestaat, wat deze inhouden en hoe ik deze competentie heb aangetoond.

De definities zijn afkomstig van het document: Business Data Management Rubrics B-competenties (Hogeschool Leiden, z.d.).

9.1.2.1 Organisatieprocessen analyseren – opdracht in context

Tabel 11: aantonen b-competentie opdracht in context

Onderdeel met definitie

Aangetoond door

Organisatiekunde – Hierbij wordt er gekeken naar de organisatie zelf. Dus wat speelt er in de organisatie, hoe zien bedrijfsprocessen eruit, wie zijn de stakeholders en wat zijn de requirements waaraan de opdracht moet voldoen, wat is de structuur en missie/visie van de organisatie

Data architectuur – Dit gaat in op de huidige data architectuur. Hierbij moeten datamodellen achterhaald/gemodelleerd worden.

Data management – Functionele gebieden van data management worden geanalyseerd. Er wordt gekeken naar de huidige situatie voor deze functionele gebieden

Eisen en wensen – Hierbij moeten de eisen en wensen geformuleerd worden. Deze moeten afgestemd worden met de stakeholders. Uiteindelijk moeten er middels methoden/technieken onderzoeksvragen tot stand komen en deze onderzoeksvragen moeten SMART zijn.

Om context aan de opdracht te geven en de lezer kennis te laten maken met het bedrijf heb ik heel veel moeten kijken naar Rijkswaterstaat zelf. Ik heb requirements gedefinieerd aan de hand van MoSCoW analyse en L-Matrix([hoofdstuk 7.2](#) en [hoofdstuk 8.2.2.4](#)), gekeken naar het proces m.b.t. de projecten en dit weergegeven in een BPMN model ([hoofdstuk 4.1](#)) gekeken naar de organisatie zelf. Dus hoe ziet de structuur er uit, waar zit ik in de organisatie, wat doet het en wat is de missie/visie ([hoofdstuk 2.1](#)). Stakeholders heb ik ook middels een stakeholders analyse in kaart gebracht ([hoofdstuk 2.3: Stakeholders](#)).

Ik heb ook gekeken naar de totale architectuur landschap binnen Rijkswaterstaaten ben ik nagegaan op welk niveau de datamodellen zich bevinden. Ik heb voornamelijk ingezoomd op de data architectuur. Middels documentatie vanuit Rijkswaterstaat heb ik deze weten te achterhalen. Ook heb ik gesprekken met data architecten gehad die aan de documentatie nog meer context gaven. [Hoofdstuk 4.2](#) legt dit allemaal uit.

Voor het advies heb ik onderzoek gedaan naar gebieden binnen datamanagement. Dit heb ik gedaan door te kijken naar twee framework en ben ik nagegaan welk framework het beste binnen de organisatie zelf past. In het advies komt dit allemaal naar voren. Zie hiervoor [hoofdstuk 8.2](#)

Middels interviews ben ik nagegaan wat de eisen en wensen waren bij de stakeholders. Deze eisen en wensen heb ik vervolgens middels MoSCoW geprioriteerd. Ook heb ik onderzoek verricht waarbij ik de onderzoeksvragen zo SMART mogelijk heb gemaakt.

9.1.2.2 Organisatieprocessen adviseren – adviseren dataoplossing

Tabel 12: aantonen b-compentie adviseren dataoplossing

Onderdeel met definitie	Aangetoond door
Context – Er wordt context gegeven aan het advies, en dit advies wordt ook in context van de organisatie geplaatst.	Ik heb uitgelegd waarom ik op deze zaken advies geef, en hierbij heb ik ook gekeken welke opties het beste past binnen de requirements gesteld door de literatuur als door Rijkswaterstaat zelf.
Advies – Hierbij wordt er een advies gegeven. Hierbij komen er verschillende oplossingen aanbod, wordt en wordt er vervolgens aangegeven welke oplossing de voorkeur heeft. Uiteindelijk moet dit afgestemd worden met de stakeholders en draagvlak binnen de organisatie gecreëerd worden.	Ik heb advies gegeven op twee gebieden waarbij er steeds twee opties waren. Deze opties heb ik uitgelegd, voor- en nadelen op een rij gezet, gekeken in welke mate dit past bij de requirements en op basis daarvan heb ik een aanbeveling gedaan voor één van de twee opties. De datamodellen heb ik ook met stakeholders besproken. Voor het advies zie hoofdstuk 8.2
Onderbouwing – Alle aangeboden oplossingen in het advies moeten onderbouwd worden, en er moet een onderbouwing komen waarom een oplossing de voorkeur heeft. Voor deze onderbouwing moet gebruik worden gemaakt van interne en externe literatuurbronnen.	In het advies heb ik gekeken naar een alternatief voor de datamodellen en naar een framework dat het beste binnen Rijkswaterstaat past. Ik heb voor beide opties voor datamodellen de voor- nadelen benoemd en uiteindelijk heb ik ook beargumenteerd waarom er voor één van de twee ontwerpen is gekozen. Dit heb ik ook gedaan voor de frameworks. Hierbij heb ik

	gekeken naar de voor- en nadelen, en weergegeven hoe goed een framework voldoet aan de eisen van Rijkswaterstaat.
Methoden en technieken – Dit advies heeft betrekking op gebied van data, het heeft dus betrekking op een data gerelateerde strategie, data architectuur, vervolgstappen voor data analyse of maatregelen op functionele gebieden van data management.	Het advies dat ik heb gegeven heeft betrekking op data management. Waarbij er naar twee frameworks werd gekeken en uiteindelijk is één framework aanbevolen.

9.1.2.3 Infrastructuur analyseren – analyseren data architectuur

Tabel 13: Aantonen b-competentie analyseren data architectuur

Onderdeel met definitie	Aangetoond door
Data architectuur – De huidige architectuur wordt geanalyseerd, het wordt gemodelleerd, in detail beschreven. Hierbij worden ook knelpunten benoemd die er eventueel zijn binnen de data architectuur. Daarnaast wordt er een oplossing of verbetering geschetst en relateert dit aan de actuele stand van zaken en technologische ontwikkelingen.	Ik heb de huidige architectuur beschreven en deze ook verder toegelicht. De lagen in de architectuur zijn visueel weergegeven. Ook zijn er minpunten benoemd aan de data architectuur. Zie hiervoor hoofdstuk 4.2 .
Datamodelleren – Datamodellen worden achterhaalt of gemaakt, hierbij wordt ervoor gezorgd dat het leesbaar en begrijpelijk is voor de stakeholders. Ook worden er knelpunten in de datamodellen benoemd en oplossingen/verbeteringen geïdentificeerd.	De opdracht stond grotendeel in het teken van data modelleren. Deze datamodellen bevatten definities zodat het voor iedereen duidelijk is wat er mee bedoeld wordt. Ik heb zelf dus nieuwe datamodellen gemaakt, maar ik heb ook een alternatief gegeven voor hoe het eventueel anders kon. Zowel de voor- als de nadelen van beide opties zijn benoemd. Dit alles is terug te vinden in hoofdstuk 8.1 .
Data management – analyseert de huidige situatie voor ieder gekozen functioneel gebied van data management. Er worden functionele gebieden van data management die betrekking hebben op de opdracht gekozen en onderbouwd. Vervolgens wordt dit geanalyseerd en mogelijke verbeteringen/oplossingen geschetst. Eventueel wordt er ook een meetinstrument ontwikkeld	Middels het onderzoek ben ik ingegaan op data management. Hierbij zijn naar twee dingen gekeken; data governance en data architectuur. Dit heb ik voor twee frameworks gedaan, DAMA DMBOK en het Cloud Data Management Capabilities (CMDCC) Framework. Uiteindelijk heb ik dan ook één framework aanbevolen. Zie hiervoor hoofdstuk 7.5 , hoofdstuk 7.6 , en hoofdstuk 8.2.2
Eisen en wensen – hierbij moet er ingegaan worden op de vraag of de huidige architectuur voldoet voor de organisatie. Er wordt onderzocht (gap analyse) in welke mate data architectuur (inclusief datamodellen) en data management in lijn zijn met de eisen en wensen binnen de organisatie. De analyse wordt vervolgens in verband gebracht met zowel de strategie van de organisatie als technologische ontwikkelingen.	In deelvraag 3 van het onderzoek heb ik gekeken naar wat de organisatie zelf al heeft gedaan aan dit probleem. Hierbij heb ik een gap-analyse uitgevoerd aan de hand van de SERVQUAL methode en ben ik nagegaan in hoeverre de organisatie voldoet aan de verwachtingen met betrekking tot de datamodellen. Zie hiervoor hoofdstuk 7.3

9.2 Planning

Situatie: Aan het begin van de stage bij het schrijven van het afstudeervoorstel heb ik een planning gemaakt. Hierbij heb ik gekeken naar alle activiteiten die ik gedurende de stage zou moeten uitvoeren

en een inschatting gemaakt van hoelang zo iets zou duren. Echter ben ik gedurende mijn stage afgeweken van de gemaakte planning. Voor de planning zie [bijlage F: afstudeervoorstel](#).

Taak: Dat er van de planning is afgeweken komt voornamelijk omdat ik had verwacht dat ik mij voornamelijk alleen bezig zou houden met het modelleren van de datamodellen. Echter is er tijdens de stage het een en ander bijgekomen. Naast de datamodellen modelleren heb ik ook advies gegeven op het onderhouden van deze modellen. Hier was bij het opstellen van de planning geen rekening mee gehouden, waardoor de planning is veranderd.

Actie: Om alsnog ervoor te zorgen dat de gewenste beroepsproducten op tijd af zijn gekomen zijn activiteiten waar in eerste instantie een langere periode voor geplant stond, sneller gedaan. Ook waren er soms momenten dat ik afhankelijk was van mijn stakeholders en de activiteit even stil stond. Op zulke momenten ging ik verder met bijvoorbeeld literatuuronderzoek. Hierdoor liepen activiteiten ook door elkaar heen.

Resultaat: Uiteindelijk zijn de beoogde beroepsproducten binnen het tijdsframe tot stand gekomen. Ondanks dat in eerste instantie bepaalde activiteiten niet mee zijn genomen in de planning is het dus toch prima gegaan.

Reflectie: Persoonlijk ben ik van mening dat het verloop van de stage goed ging. Ik had geen problemen met tijdnoed en mocht het zo zijn dat ik ergens even niet verder kon, dan waren er altijd wel andere activiteiten die ik kon oppakken. Een eventuele verbetering voor in de toekomst met betrekking tot plannen, is om meer tijd te reserveren voor 'onverwachte' zaken. Dit had ik tijdens deze stage niet gedaan, maar omdat ik voor sommige activiteiten meer tijd had ingeschat en dit achteraf sneller verliep dan in eerste instantie verwacht kwam ik niet in tijdnoed. Ik kan mij voorstellen dat wanneer dit niet het geval was, ik zeker weten in tijdnoed kwam en mogelijk niet alle producten af kon krijgen.

9.3 Uitkomsten

Situatie: Voor de stage opdracht heb ik datamodellen gemaakt en uiteindelijk een advies gegeven. Deze zijn aan het einde van de stage opgeleverd.

Taak: Om uiteindelijk tot een goed resultaat te komen heb ik een onderzoek opgesteld. Elke deelvraag draagt bij aan het beantwoorden van de hoofdvraag. Zodra de hoofdvraag dan 'beantwoord' was, leidde dit tot de beroepsproducten.

Actie: Hiervoor heb ik onderzoek gedaan naar relevante onderwerpen die hielpen om zo correct mogelijke producten tot stand te brengen. Hiervoor zijn verschillende onderzoeksmethodes voor gebruikt en uiteindelijk zijn alle deelvragen beantwoord.

Resultaat: Uiteindelijk zijn er datamodellen tot stand gekomen die ook in de 'echte' omgeving van Rijkswaterstaat staan en is er een advies gekomen over hoe ze eventueel deze datamodellen kunnen onderhouden

Reflectie: Ik ben van mening dat de beroepsproducten prima zijn. Alles is binnen het tijdframe opgeleverd. Bij het uitvoeren van de stage heb ik ook veel geleerd. Zowel over vakinhoudelijke dingen als algemene zaken.

9.4 Risico's

Bij het schrijven van het afstudeerplan zijn er zes scenario's geschetst die mogelijk een risico vormde voor het verloop van de stage. Vier van de zes risico's hebben zich voorgedaan. De volgende risico's (benoemd in het afstudeerplan) speelde op tijdens de stage:

1. Wanneer bepaalde zaken in het project geblokkeerd zijn. Sparx heeft een functie waarbij alleen bepaalde mensen de datamodellen kunnen bewerken
2. Een ander risico is wanneer de stakeholder moeilijk te bereiken zijn. De stakeholder hebben veel meer verstand van de data zelf, en hun input zal geregeld nodig zijn om de data beter te begrijpen. Het kan lang duren voor dat ze antwoorden

3. Wanneer verschillende mensen een andere definitie hebben voor dezelfde entiteit. Het is een grote organisatie met veel medewerkers
4. Momenteel worden de documenten met de conceptuele modellen nog steeds bijgewerkt. Drastische aanpassingen aan de documenten zal ervoor zorgen dat ook de UML-modellen aangepast moeten worden.

Toen risico één zich voordeed heb ik contact opgenomen met de persoon die verantwoordelijk is voor het toekennen van bevoegdheden. Uiteindelijk ging dit nog niet helemaal vlekkeloos en heeft dit toch wel een paar weken in beslag genomen om dit weer goed te krijgen.

Bij risico twee heb ik voornamelijk gekeken naar hoeveel meerwaarde de antwoorden van zo'n persoon is. Als er dus meerdere stakeholders waren die dezelfde functie bekleden en hoogstwaarschijnlijk hetzelfde antwoord gaven dan was het niet zo heel cruciaal om een gesprek met die persoon te voeren. Mocht het wel zo zijn dat ik echt van die persoon een antwoord nodig had dan stuurde ik nog een mail. Wanneer daar ook niet op gereageerd werd heb ik gekeken of er andere mensen waren die mij daarbij konden. Deze situatie heeft zich een keer voorgedaan en toen kon iemand anders hier mij ook meehelpen.

Voor het derde risico ben ik steeds naar mijn stagebegeleider gegaan en dit bij hem aangekaart. Ik heb verteld wat volgens de andere stakeholder de juiste definitie was. Uiteindelijk ging mijn stagebegeleider hier ook verder naar kijken en hierover met andere mensen in gesprek. Wanneer ze toe de conclusie waren gekomen dat de definitie zo wel goed was heb ik dit aangepast in het model. Om risico vier te voorkomen heb ik steeds om de paar weken het Word document vergeleken met de modellen. Ik ben steeds alles nagegaan en wanneer er iets veranderde heb ik dit ook in de modellen aangepast.

Terugkijkend op dit alles was het toch handig dat ik van te voren al scenario's had geschetst waarbij risico's zich konden voordoen. Hierdoor wist ik in principe al wat ik kon doen als het probleem zich voordeed. Hier heb ik dus ook van geleerd dat het altijd handig is om vooraf een risicoanalyse te doen, waardoor je niet opeens voor hele onverwachte dingen komt te staan. Door eerst al na te denken over risico's en mogelijke maatregelen die dan genomen kunnen worden, ben je al voorbereid op zulke dingen.

9.5 Begeleiding

Tijdens het pitchen van de stageopdracht kwam naar voren dat de opdracht mogelijk niet complex genoeg was. Het zou een brede opdracht zijn maar niet moeilijk genoeg omdat het alleen over datamodellen ging. Dit heeft gedurende de stage mij wel redelijk wat stress gegeven. Door in goed gesprek te gaan met mijn begeleider vanuit school kon dit probleem worden opgelost. Er is gekozen om uiteindelijk ook een advies eraan toe te voegen, om het toch meer diepgang te geven. Ook had ik soms wat vragen omtrent de competenties. Ook daarmee kon hij mij altijd helpen.

Ook de begeleiding vanuit het stagebedrijf was goed. Er waren wekelijkse gesprekken waarbij de voortgang besproken werd. Bij vragen kon ik hem altijd mailen of een bericht op Microsoft-teams sturen. Hij wist dan altijd wel het antwoord en anders wist hij wel wie mogelijk een antwoord voor mij had of had hij documenten die mij konden helpen. De begeleider zorgde er ook voor dat ik de juiste mensen vond die mij verder konden helpen.

Kortom, zowel de begeleiding vanuit school als vanuit het stagebedrijf waren uitstekend. Ik kon altijd bij ze terecht en vragen werden altijd beantwoord.

10. BIJLAGE

10.1 Bijlage A: Literatuurlijst

10.1.1 Afbeeldingen

Booch, G., & Rumbaugh, J. (2005a). Association Names [Illustratie]. In *The Unified Modeling Language User Guide* (2de editie, p. 63)

Booch, G., & Rumbaugh, J. (2005b). Roles [Illustratie]. In *The Unified Modeling Language User Guide* (2de editie, p. 63)

Booch, G., & Rumbaugh, J. (2005c). Dependencies [Illustratie]. In *The Unified Modeling Language User Guide* (2de editie, p. 61).

Booch, G., & Rumbaugh, J. (2005d). Generalization [Illustratie]. In *The Unified Modeling Language User Guide* (2de editie, p. 62).

DMBOK2. (2017). Architecture domains [Tabel]. In *DAMA-DMBOK Data management body of knowledge* (2de editie, pp. 103–104).

G'Enova, G., Llorens, J., & Martí iNez, P. (2002, 2 december). *A classical example of binary association with the expression of multiplicities* [Illustratie]. Special Issue UML 2001.
<https://link.springer.com/content/pdf/10.1007/s10270-002-0009-3.pdf>

Kamp, M. W. (2009). Bedrijfsobjectmodel 2009 [Objectmodel]. In *Prolaborate Rijkswaterstaat* (p. 1).

Rijkswaterstaat. (2021b, januari 28). *Data-architectuur: Samenhang* [Illustratie]. Rijkswaterstaat Doelarchitectuur Data. <https://werkwijzer.apps.cf-am2.intranet.rws.nl/index2.html>

Rijkswaterstaat. (2021c, mei 31). *Organisatieonderdelen* [Organogram]. Rijkswaterstaat.
<https://corporate.intranet.rws.nl/Organisatie/Organisatieonderdelen/>

Rijkswaterstaat. (2022, 14 februari). *Centrale Informatievoorziening* [Organogram]. Rijkswaterstaat.
https://corporate.intranet.rws.nl/Organisatie/Organisatieonderdelen/Centrale_Informatievoorziening/

Statista Research Department. (2022, 23 mei). *Volume of data/information created, captured, copied, and consumed worldwide from 2010 to 2025* [Grafiek]. Statista.
<https://www.statista.com/statistics/871513/worldwide-data-created/>

10.1.2 Websites

Bellekens, G. (2018, 23 april). UML Best Practice: 5 rules for better UML diagrams. Bellekens. Geraadpleegd op 17 mei 2022, van https://bellekens.com/2012/02/21/uml-best-practice-5-rules-for-better-uml-diagrams/?gclid=EAIaIQobChMIsJPd8tOJ-AIVBeh3Ch3DOgd_EAAYBCAAEgJvffD_BwE

Hogeschool Leiden. (z.d.). Business Data Management Rubrics B-competenties. Hogeschool Leiden sharepoint. Geraadpleegd op 15 december 2021, van <https://hogeschoolleiden.sharepoint.com/sites/og2122.iwlab->

afstuderen/Moduledocumentatie/Forms/AllItems.aspx?id=%2Fsites%2Fog2122%2Eiwlab%2Dafstuden%2FModuledocumentatie%2FRubrics%2FBijlage%2006%20%2D%20Rubrics%20B%2Dcompetenties%20BDaM%2Epdf&parent=%2Fsites%2Fog2122%2Eiwlab%2Dafstuden%2FModuledocumentatie%2FRubrics

Hogeschool Leiden. (2021, 20 augustus). *Informatica Rubrics A-competenties* [Presentatieslides]. SharePoint. <https://hogeschoolleiden.sharepoint.com/sites/og2122.iwlab-afstuderen/Moduledocumentatie/Forms/AllItems.aspx?id=%2Fsites%2Fog2122%2Eiwlab%2Dafstuden%2FModuledocumentatie%2FRubrics%2FBijlage%20N%20%2D%20Rubrics%20A%2Dcompetenties%2Epdf&parent=%2Fsites%2Fog2122%2Eiwlab%2Dafstuden%2FModuledocumentatie%2FRubrics>

managementmodellensite. (2021, 21 maart). *L-matrix*. Geraadpleegd op 14 juni 2022, van <https://managementmodellensite.nl/l-matrix/>

Ministerie van Infrastructuur en Waterstaat. (2021a, april 16). Rijkswaterstaat Centrale Informatievoorziening. Rijkswaterstaat. Geraadpleegd op 21 december 2021, van <https://www.rijkswaterstaat.nl/over-ons/onze-organisatie/organisatiestructuur/centrale-informatievoorziening>

Ministerie van Infrastructuur en Waterstaat. (2021b, december 13). Onze organisatie. Rijkswaterstaat. Geraadpleegd op 21 december 2021, van <https://www.rijkswaterstaat.nl/over-ons/onze-organisatie>

Ministerie van Infrastructuur en Waterstaat. (2021c, 20 april). Missie. Rijkswaterstaat. Geraadpleegd op 13 juni 2022, van <https://www.rijkswaterstaat.nl/over-ons/onze-organisatie/onze-missie>

Planview projectplace. (2021, 25 juni). *Projectplace – Project Management Tools and Features*. Planview. Geraadpleegd op 14 mei 2022, van <https://www.planview.com/products-solutions/products/projectplace/projectplace-features/>

Rijkswaterstaat. (2020, 16 november). *Afdeling Strategie & Beleid*. Intranet Rijkswaterstaat. Geraadpleegd op 18 april 2022, van https://corporate.intranet.rws.nl/Organisatie/Organisatieonderdelen/Centrale_Informatievoorziening/Directies_en_afdelingen/Bedrijfsvoering_Inkoop_en_Strategie_BIS/Afdeling_Strategie_Beleid/

Sparx Systems. (2022a). Full Lifecycle Modeling for Business, Software and Systems | Sparx Systems. Geraadpleegd op 25 mei 2022, van <https://sparxsystems.com/products/ea/>

Steenbeek, I. (2021, 12 april). *DAMA-DMBOK2 vs DCAM®2.2: Usage Statistics*. Data Crossroads. Geraadpleegd op 14 juni 2022, van <https://datacrossroads.nl/2021/03/31/dama-dmbok2-vs-dcam-2-2-usage-statistics/>

10.1.3 Boeken

Booch, G., & Rumbaugh, J. (2005b). *Unified Modeling Language User Guide, The (2nd edition)*. Pearson Education (US).

Dama International, Dama International, Data Management Association, Henderson, D., & Earley, S. (2017). *DAMA-DMBOK*. Technics Publications.

Desfray, P., & Raymond, G. (2014). *Modeling Enterprise Architecture with TOGAF* [E-book]. Elsevier Gezondheidszorg.

EDM Council Inc. (2021). *Cloud Data Management Capabilities Framework* (1.1.1 ed.). EDM Council Inc. https://cdn.ymaws.com/edmcouncil.org/resource/resmgr/cdmc_master/CDMC_Framework_V1.1.1.pdf

10.1.4 PDF-documenten

IBM. (2015, augustus). UML basics: The class diagram. https://www.softwareresearch.net/fileadmin/src/docs/teaching/WS13/SE/UML_basics-The_class_diagram.pdf

Kamp, M. (2021, september). Databeheerprincipes. Rijkswaterstaat.

Loonen, T. (2010). De elementen van een klassediagram. *Database magazin*, 6, 26–29. <https://biplatform.nl/magazines/Aveq/117898.pdf>

Rijkswaterstaat. (2018, oktober). *RIVA 2022*.

Rijkswaterstaat. (2021a, januari). *Rijkswaterstaat Doelarchitectuur Data*. <https://werkwijzer.apps.cf-am2.intranet.rws.nl/index2.html>

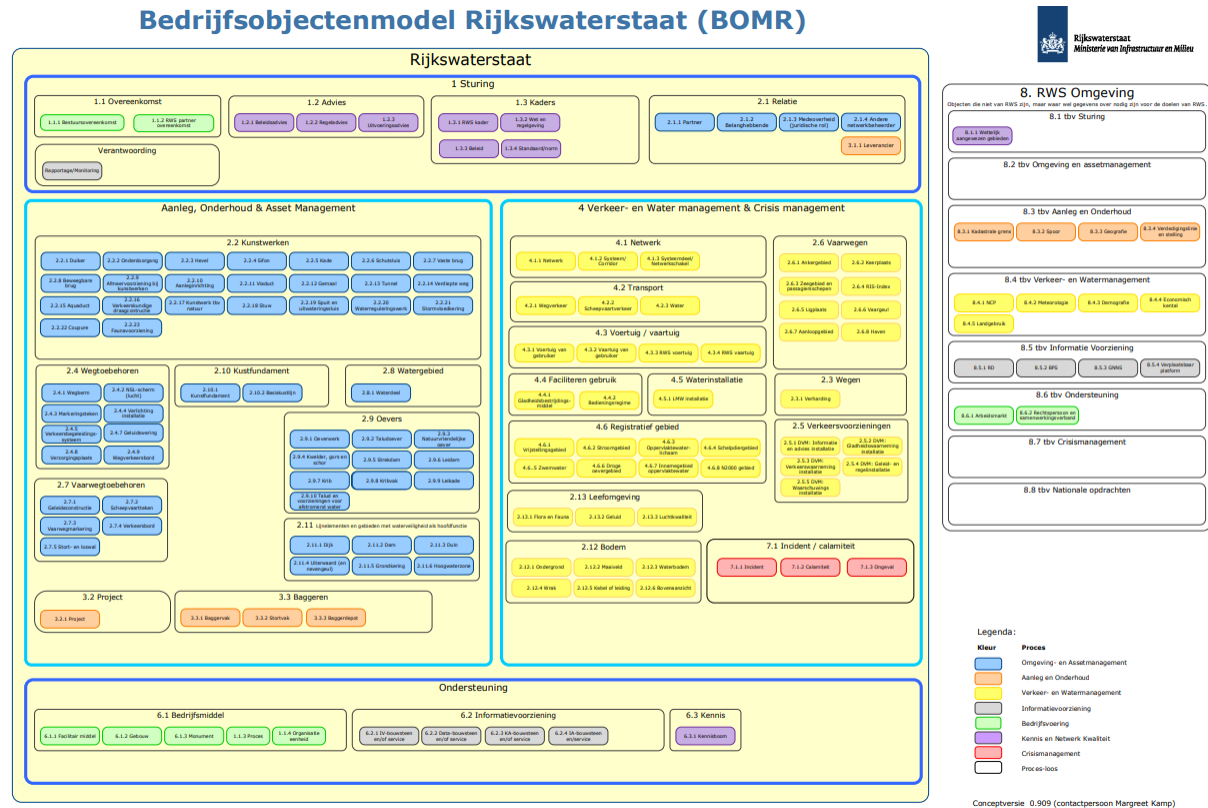
Rijkswaterstaat. (2021d). Een sterke Informatievoorzieningsfunctie, een flexibele organisatie.

Sparx Systems. (2022b). *hich Enterprise Architect Edition Should I Purchase?* <https://sparxsystems.com/downloads/pdf/editions.pdf>

Wilkinson, M. D. (2016). The FAIR Guiding Principles for scientific data management and stewardship. *Scientific data*, 1–4. <https://www.nature.com/articles/sdata201618.pdf>

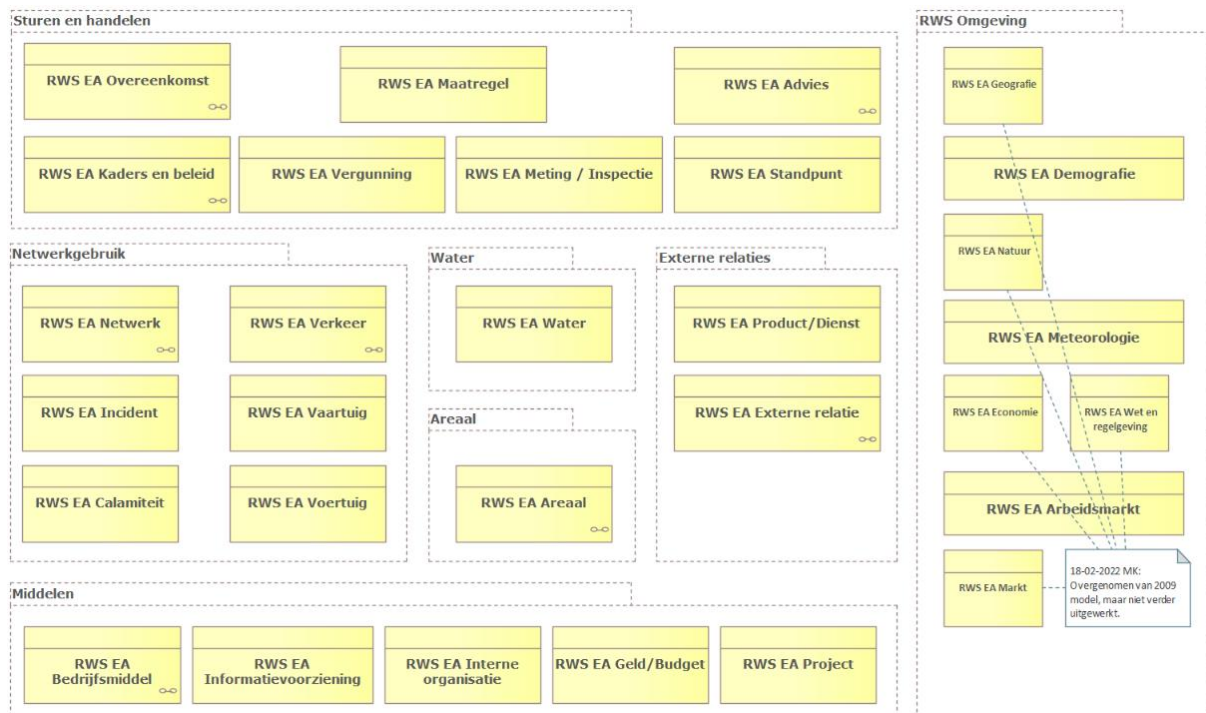
10.2 Bijlage B: Het bedrijfsobjectmodel RWS (BOMR)

10.2.1 Het huidige bedrijfsobjectmodel RWS (BOMR)



10.2.2 Het bedrijfsobjectmodel RWS (BOMR) 2009

Onderstaand model is de het BOMR uit 2009 maar dit is opnieuw geïntroduceerd en dient als uitgangspunt. Er komt namelijk een nieuwere versie hiervan waarbij dit iets meer in detail treedt zodat het voor de organisatie beter bruikbaar wordt. Deze is momenteel nog niet aanwezig.



Figuur 19: Objectmodel RWS (Kamp, 2009)

10.3 Bijlage C: Interviews

10.3.1 Interview/vragenlijst Victor van der Kloet

Ingevuld door: Victor van der Kloet

Ingevuld op: 30-3-2022

Zou ze jezelf misschien kunnen voorstellen (naam en functie)?

Ik ben Victor van der Kloet, informatieanalist bij de afdeling IV Platformen binnen de directie IRN van de CIV van Rijkswaterstaat.

Hoe gaat Rijkswaterstaat om met de data voordat er datamodellen waren?

Data valt onder de afdeling Data Management Center (DMC) binnen de directie IGA van de CIV van Rijkswaterstaat. Data architect Margreet Kamp kan hier meer over zeggen. Bij IVP beheren we alleen het datamodel van de CMDB, maar dit hoort eigenlijk ook bij het DMC te liggen.

In de opdracht kwam naar voren dat er ook definitieproblemen zijn en daarom dat er nu datamodellen vastgelegd worden, hoe is dit verledentijd opgelost/aangepakt?

Er is binnen RWS geen centraal datamodel. Hier wordt binnen verschillende domeinen decentraal in voorzien. De CD heeft bijvoorbeeld een datamodel voor het HR-domein ontwikkeld. En zie ook het reeds genoemde datamodel van de CMDB.

Hoe beheert/onderhoudt Rijkswaterstaat nu de datamodellen? Hoe bewaakt Rijkswaterstaat de data kwaliteit?

Beheer datamodellen: zie hiervoor, er is geen centraal beheer van datamodellen.

Datakwaliteit: hiervoor is een datakwaliteitsstelsel ontwikkeld, zie op het intranet

https://corporate.intranet.rws.nl/Kennis_en_Expertise/Data/Hoe_is_de_datakwaliteit_geregeld/Kader_Datakwaliteitsstelsel/.

Wat gebeurt er momenteel aan het modelleren van de datamodellen?

Ik heb hier niet het complete beeld van. Naast het datamodel van de CMDB (inmiddels eigenlijk IV Services en IV Assets) heb ik met de opsteller van deze vragenlijst gewerkt aan de ontwikkeling van het datamodel voor IV-projecten als onderdeel van een datamodel Bedrijfsvoering IV in het kader van het programma Datagedreven Sturen.

Is er vanuit Rijkswaterstaat een standaard waar de datamodellen aan moeten voldoen?

De datamodellering gebeurt volgens UML, m.b.v. Sparx Enterprise Architect. Ik heb geen weet van een verdergaande standaard behalve aansluiten op het Bedrijfsobjectenmodel Rijkswaterstaat (BOMR) als bovenliggend architectuurmodel.

Zijn er vanuit RWS requirements waar de datamodellen voor iv-projecten aan moeten voldoen?

Functionele en non-functionele requirements?

Het datamodel voor IV-projecten in Sparx EA is een UML-vertaling van een schriftelijke beschrijving in ERD-terminologie die in het kader van het programma Datagedreven Sturen door een werkgroep onder leiding van Rob Smits is opgesteld. Het moet dus een correcte weergave in UML van dit schriftelijke datamodel zijn.

In hoeverre is de (data) architectuur en data management in lijn met de eisen en wensen van Rijkswaterstaat?

Dat is een goede vraag, maar voor een antwoord wil ik verwijzen naar de data architect Margreet Kamp.

In hoeverre wordt er gekeken naar de UML principes bij het modelleren van de data?

Binnen het tool Sparx EA wordt gebruik gemaakt van de UML-ondersteuning die dit biedt en in de training voor het werken hiermee wordt één dag besteed aan het modelleren volgens de principes van UML.

Wordt er binnen de gehele organisatie datamodellen vastgelegd of verschilt dat per afdeling?

Zoals hiervoor aangegeven verschilt dit per afdeling.

Waarom is er gekozen om Sparx te gebruiken voor het modelleren van de datamodellen?

Vanuit de informatiemanagers en architecten is op een gegeven moment, na een vraag hierover van informatiemanager Rob Smits aan architect Peter Bernhard, gekozen voor Sparx. Waarom dat weet ik niet precies meer, misschien kan collega en voormalig informatiemanager Mieke van Nierop of Renée Mantel (applicatiebeheerder) deze vraag beter beantwoorden.

Gekeken naar de 11 kennisgebieden van DAMA DMBOK, welke gebieden zijn in jouw optiek handig om rekening mee te houden? En waarom?

Volgens mij moet je altijd al deze kennisgebieden goed willen afdekken, maar als ik kijk naar mijn nieuwe werkgever, gemeente Rotterdam, waar ik de rol van lead data architect ga vervullen, dan ligt daar nu vooral veel nadruk op data integration and interoperability en op metadata.

10.3.2 Interview Peter Bernhard

Datum van interview: 13-04-2022

Geïnterviewde: Peter Bernhard

Geïnterviewt door: Bo-Lan Schouten

Spreker 1 Bo-Lan Schouten

Spreker 2: Peter Bernhard

Spreker 1: Mijn eerste vraag was of je jezelf misschien kan voorstellen, wie je bent, je functie?

Spreker 2: Okay Peter Bernard domein architect watermanagement bij de Rijkswaterstaat CIV

Spreker 1: En ik vroeg mij ook af of je misschien weet hoe Rijkswaterstaat omgaat met de data, voordat er echt datamodellen waren, want ik heb begrepen dat jullie eerst nog geen datamodellen hadden eigenlijk en dat jullie daar nu een beetje mee bezig zijn om dat te maken.

Spreker 2: ja hoe gaat Rijkswaterstaat om met zijn data, dat is natuurlijk een lastige vraag. Rijkswaterstaat is een wil erg veel koppige organisatie die eigenlijk veel verschillende dingen doet, neem domein watermanagement. Die hebben een organisatie opgericht, die heet informatie uit water en die hebben een datamodel opgesteld voor het Nederlandse water in het datamodel dat is dan van toepassing op alle waterbeheerders in Nederland. Bij scheepvaartmanagement hebben ze een soortgelijke exercitie gedaan maar dan veel Europeser. Dus scheepvaart gaat natuurlijk heel snel de grens over dus dan moet het datamodel ook breder worden afgestemd dus collega's van mij zitten ook regelmatig in wenen om daar over datamodellen nou te spreken en verkeersmanagement en zo heeft ook zo z'n eigen organisatie om datamodellen af te spreken dus als je vraagt hoe gaat Rijkswaterstaat om met z'n data en datamodellen? Nou sommige doen het heel erg goed alleen heeft eigenlijk met Rijkswaterstaat als sturende factor niet zoveel van doen omdat al die datamodellen al heel snel zich oriënteren op de rest van de community buiten Rijkswaterstaat

Spreker1: Okay, maar ik mij ook afvroeg dan, want er waren eerst geen datamodellen voor zover ik heb begrepen, hoe gingen jullie toen om met de data?

Spreker 2: Het is heel simpel als het in de database zit is de data.

Spreker 1: oké en jullie hadden toen nog geen behoefte aan datamodellen?

Spreker 2: Ja die behoefte is er natuurlijk altijd wel geweest alleen het is vrij lastig om alles te organiseren en de uitspraak we hebben geen datamodellen klopt ook niet helemaal, dus er zijn wel degelijk datamodellen alleen die zitten heel snel in organisaties buiten Rijkswaterstaat dus, in die zin heb je gelijk Rijkswaterstaat heeft geen datum modellen dus niet eigen datamodellen. Aders komt Rijkswaterstaat ook op een eiland te staan, maar we gebruiken ze wel degelijk.

Spreker 1: Dan is het van anderen bedrijven ofzo externen bronnen?

Spreker 2: Ja voorname vanuit meer internationale of nationale organisaties waarin we die datamodellen afspreken en soms is Rijkswaterstaat zo'n krachtige organisatie dat we het voortouw nemen in dat soort datamodellen bijvoorbeeld bij aanleg en onderhoud is Rijkswaterstaat wel één van de sturende partijen in het datamodel over assets en dingen van beton en staal die wij omdat wij n eenmaal de grootste opdrachtgevers zijn voor dit soort kunstwerken en aannemer klussen, dus daarin neemt Rijkswaterstaat dus ook vaak het voortouw in het opstellen van dit soort datamodellen.

Spreker 1: Goed om te horen. Mijn volgende vraag is: ik had dus een opdracht gekregen vanuit Rijkswaterstaat om natuurlijk datamodellen in Sparx te zetten en dat was vooral, van wat ik had begrepen doel daarvan, om ook definities vast te leggen. Mijn vraag was hoe jullie de definitieproblemen verledentijd hebben opgelost?

Spreker 2: Deze vind ik moeilijk om te beantwoorden want we hadden niet heel veel last van definitieproblemen. Meestal was het zo op voorhand wel eenduidig wat wat is. Eén kilometer asfalt is een kilometer asfalt en waterhoogte is een waterhoogte. Er zaten wel wat dingen in definitie vraagstukken over de wat meer specifieke dingen, zeker als je biologie gaat meten in water en je gaat je afvragen, is dat nou gevaarlijk, is dat giftig, is dat voldoende? Dan moet je wel goed kunnen definiëren wat is de maximale hoeveelheid van een stof die in water mag zitten. Is dat per kubieke meter of is dat per mol is dat per liter of is dat per stroom eenheid? Daar zaten nog wel wat verschil in, maar meestal werd dat wel pragmatisch opgelost.

Spreker 1: nee want Rob [Smits] heeft tegen mij gezegd dat er vooral problemen waren met bijvoorbeeld de projecten. Je hebt een project en het project gaat over om van de ene kant van een rivier naar de andere kant van de rivier te komen maar dit kan dus een tunnel zijn maar dit kan ook een brug zijn. En dan kunnen het verschillende soorten bruggen zijn. Dat daar een beetje het probleem was van de definitie. Dit kan soms onduidelijk zijn of bijvoorbeeld als je kijkt naar een startdatum. Wanneer is de startdatum? Is dat het moment dat de opdrachtgever heeft gezegd ik heb een probleem, kunnen jullie daar een oplossing voor vinden. Of wanneer de projectleider een aanvraag of iets dergelijks ingediend heeft.

Spreker 2: Ja precies in het domein waarin je werkt, als je bijvoorbeeld zegt we moeten over een rivier en daar moet de brug komen en dan welk soort brug dan precies en hang een ding of een staande ding, binnen dat domein zelf zijn daar wel scherpe definities van. Dat is het probleem niet, het probleem ontstaat vooral en dat is natuurlijk wel het probleem van Rijkswaterstaat dat zo'n vormig monster is, de vraag is snapt de rest van Rijkswaterstaat dat dan ook? De grap is, is de brug open of dicht? Dat hangt ervan af. Als de brug dus in de verticale toestand staat dan is hij open voor het scheepvaartverkeer, dus vanuit scheepvaart zeggen ze de brug is open. Vanuit het land verkeer zeggen we, nee de brug is dicht want ik kan er met mijn auto niet overheen. Dus de definitie van wat is een brug die open is, daar zijn we volgens mij nog steeds niet over uit. Dan zie je dus dat de definitie kwesties zitten met name op dat domein overstijgende stuk, dus de vraag is hoe hoog is de brug? Nou ja vanuit het brugbeheer, daar kijken ze in feite van: de brug ten opzichte van het nationaal Amsterdams peil. Of hoe hoog is de brug ten opzichte van het land? Dan zien we dus, de brug is drie meter hoog ten opzichte van het land dus je hebt een talud nodig dat het verkeer drie meter hoger brengt. Dat is de hoogte van de brug wat hun betreft, maar van het scheepvaart is de hoogte van de brug afhankelijk opzichte van dat water.

Spreker 1: Maar als jullie dan zulke ja problemen hebben als waar hoe hebben jullie dat opgelost of zijn zulke dingen nog steeds onduidelijk?

Spreker 2: Ja dat betekent dat je heel veel moet vertalen, dus het uitgangspunt is dan wel hoe hoog is de brug zoals het wordt opgeleverd? Bij vaarwegmanagement maken ze dan een vertaling van hoe hoog is de brug ten opzichte van het gemiddelde peil en zo krijg je dus per domein een verschillende brughoogte die dus in verschillende communities wordt gedistribueerd. Zolang die verschillende definities in verschillende communities blijven is dat helemaal niet zo'n probleem. Zodra ze dus door de communities heen beginnen te lopen, dat ze zeggen nee je moet voor het verkeer, en voor de mensen die de brug moeten verven, en voor de mensen die onder de brug door moeten varen, dan moeten we allemaal dezelfde definitie hebben van hoe hoog is die brug. Nee want als die brug hoger is moet je meer emmers verf meenemen om die brugpijlen te verven. Daar verschijnen inderdaad de definitieproblemen. Dat is ook inderdaad een kwestie dat door heel veel nachecken van op moment dat iemand zegt van deze brugpijler is vijf meter hoog wilt u hem verven, eerste wat je moet vragen: ja vijf meter hoog ten opzichte van wat?

Spreker 1: Dat je het heel specifiek maakt eigenlijk.

Spreker 2: Ja dat moet dus iedere keer weer uiteindelijk ontstaat daar wel een werkwijze over. De definitieproblemen ontstaan dus eigenlijk vooral op de op de stukken tussen de verschillende domeinen waren rijkswaterstaat acteert en binnen zo'n domein zijn de definities over algemeen wel scherp.

Spreker 1: Hoe beheert Rijkswaterstaat nu de datamodellen en hoe bewaakt de data kwaliteit?

Spreker 2: Ja dat zijn nog twee verschillende verschillende vragen. Op papier hebben we het over het algemeen datamodel beherende organisaties die een datamodel beheren dus we hebben informatie op water die beheert aquodata model waarin beschreven staat wat wat is. Scheepvaart heeft ook zoiets en aanleg/onderhoud heeft ook zoiets. In de praktijk is het datamodel datgene wat in de database staat. Want als je iets met die data wil gaan doen, je wilt kijken van oké hoeveel water heb ik en wat is de kwaliteit daarvan? Dan kijk je vooral naar de data in de database en het datamodel wat in die database zit is dan bepalend voor wat wat is en dat is ook niet altijd in overeenstemming met het papieren model dat we afgesproken hebben. Een de facto is het datamodel, datgene dat in de database staat, zoals die in de database gedefinieerd is dat is wat het is. Daar werken we echt mee.

Spreker 1: Dat ging op papier dan anders is waar dient dat dan voor?

Spreker 2: Zodra je dus het systeem gaat verbouwen, zodra je nieuwe dingen gaat doen, dus we gaan nu met slim watermanagement onze waterdata van verschillende waterbeheerders naast elkaar leggen en zul je zien waterschap 1 heeft het zo gedaan, Rijkswaterstaat heeft het zo gedaan, en een ander waterschap heeft het weer anders gedaan, dan heb je wel degelijk een papieren model waarop je zegt ja maar zo hadden we het afgesproken. Dus leuk dat jij intern in jouw organisatie iets anders hebt ze te freubelen maar zodra we met elkaar gaan communiceren moet je dat toe praten naar het papieren model en anders accepteren we het niet en dat gebeurt ook daadwerkelijk. Op het moment dat wij dus onze waterkwaliteit naar Brussel moeten rapporteren, daar is Europese wetgeving voor, dus iedereen moet aan Brussel rapporteren wat de waterkwaliteit is conform Europese normen. Dan moeten alle water beheren dus hun water conform het afgesproken datamodel aanleveren of informatie als water gaat ze daar naartoe converteren. En dan wordt het allemaal onderling op telbaar en dan kunnen we een nationale rapportage. De spanning tussen het fenomeen datamodel en de data zoals we die feitelijk gebruiken, die wordt groter naarmate we die data breder gaan gebruiken. Dan wordt het verschil tussen een papieren datamodel en je database dat wordt steeds vervelender en dan gaan we er dan ook meer aan doen.

Spreker 1: Hoe bewaak je dan de kwaliteit?

Spreker 2: het verschilt nogal per domein. Het watermanagement domein is heel sterk centraal georganiseerd, dus we hebben een landelijk meetnet dat op landelijk niveau alle waterdata meet zowel in kwaliteit dus met laboratoria die meten hoeveel gifstoffen zitten er in het water als in kwantiteit dus hoe hoog staat het water in hard stroomt het. Dat wordt allemaal centraal neergezet en omdat het centraal wordt neergezet kunnen we ook centraal de data monitoren én kwalitatief goed maken, dus er is een afdeling bij de CIV, IGA [Inwinning en gegevens analyse] dus de directie IGA, en er zitten mensen die de data in de gaten houden, over het algemeen met de hand. We zijn nu bezig om daar dan een stukje software tegenaan te zetten die uitschieters in de metingen detecteren en zegt hé hier vliegt iets uit de bocht. Zou er een schip tegen metaal aan gevaren zijn of zou dat ergens de meettijd verstopt zitten of de sensor verlopen? Dus die houden dat in de gaten en zo kun je dus datakwaliteit in de gaten houden. Maar andere processen zijn veel meer afhankelijk van wat er

regionaal in een project wordt opgeleverd én wat voor de aannemer aan databestanden aanlevert en dan moeten we maar hopen dat die aannemer het goed gedaan heeft. Dan is het van veel meer afhankelijk van wat er dan specifiek in die regio nog gebeurt aan datakwaliteit om te kijken oké wat weten we van de datakwaliteit. Daar is de datakwaliteit eigenlijk meer diffuus. Een belangrijke beweging die dus Rijkswaterstaat heeft gestart is de data APK. Normaal stuur je een auto naar de APK om te kijken of je er mee de weg op mag. Dat willen we ook voor de data doen dus daar hebben wij tooling voor die dan een databestand langs een tool haalt en die meet dan van hé ik heb hier een uitschieter, of ik heb hier een gat in mijn data, of ik heb hier een ding dat niet binnen die range hoort te vallen; opeens een brug van tweehonderd meter hoog of zo dat kan niet kloppen en die software gaat dat er dus uit vissen. Daar hebben we dus een kwaliteitsverhogende actie gestart. Maar dat is allemaal nog onderhanden werk dus dat moet nog dit jaar worden teruggekoppeld naar de databasebeheerder. Die moet zeggen je hebt een brug staan die tweehonderd meter hoog is en ja dat klopt inderdaad niet maar dan moet het aangepast worden. Het moet weer getoetst worden, dan pas is het goed. Dat proces begint nu te lopen, dat is een vrij recent proces.

Spreker 1: Op zich wel handig dat er zoiets is eigenlijk, ondanks dat het in de beginfase is is het wel een goede check. Nog een andere vraag wat gebeurt er momenteel aan het modelleren van de datamodellen? Ik had begrepen dat ze eerst in word bestand stonden en nu worden ze een beetje visueel gemaakt, maar volgen mij was er nog niet heel veel.

Spreker 2: Nee mijn ervaring is inderdaad, ik heb in de afgelopen tien jaar geen systematische vastlegging van datamodellen gezien binnen Rijkswaterstaat.

Spreker 1: Weet je ook precies waarom dat komt waarom dat het niet eerder is vastgelegd?

Spreker 2: Omdat niemand het een probleem vond. We zijn erg druk gestuurd. Waar iemand een probleem ervaart gaat hij oplossen. Zodra iemand zit; op papier zou er een probleem moeten zijn maar ik ervaar het niet zo, dan gaat niemand wat doen. Verschuiving die je nu bij Rijkswaterstaat ziet optreden is dat steeds meer mensen door beginnen te krijgen dat zij data gedreven moeten gaan werken. Want zolang data worden vastgehouden voor een proces, je hebt bijvoorbeeld een stuk weg in beheer. Ik heb een database en er is hier een weg uitgetekend. Dus veel kilometers, zo breed na twee rijstroken parkeren zoveel lantaarnpalen. Dat is dan in de database en er is een club die onderhoudt die weg aan de hand van die data in die database. Voor de rest bemoeit ze zich niemand zich met die weg. Er gaat niemand zich in enige mate zorgen maken over een later model. Data zit toch in die database, inderdaad, is keurig onderhouden, nou de lichten doen het, dus de gaten in het asfalt zijn gevuld dus wie loopt hier nou te zeuren. Nou niemand dus niemand boeit dat totdat een minister vraagt hoeveel kilometer zoab heb ik eigenlijk. ja dan is het feit dat het allemaal in verschillende databases zit zonder bovenliggende datamodel dat is dan opeens een probleem geworden nou ja dat is de minister dat vraagt en dat is iedereen van oh ja de meeste wat vraag nou haal al die databases met elkaar stop ze in de spreadsheet tel het op trek het af, menigvuldig het, schop dat getal naar de tweede kamer en we zeggen 2316 kilometer en 50 centimeter. Klaar, volgende opdracht én de minister zegt ongeveer twintigduizend kilometer zoab en dan is het ook weer voor elkaar. Dat heeft geen impact, maar je krijgt nu systematisch dat de rapportages gevraagd moeten worden, en systematisch dat mensen zeggen nou ja als je zoveel kilometer afstand hebt staan krijg je zoveel geld én als niet dus moet wel kloppen én omdat dat soort processen beginnen te lopen zouden steeds meer mensen naar die data willen kijken, wordt het steeds vervelender dat je geen datamodel en dat je datamodel niet klopt met de datamodel dat wordt steeds vervelender, en als het vervelend wordt gaan we er wat aan doen.

Spreker 1: nee interessant, wel op zich bijzonder. Lijkt me toch wel dat het niet eerder is vastgelegd. Maar nu zijn ze bezig met het modelleren ervan, weet je ook in hoeverre dat al gaande is? Zijn ze er al redelijk ver mee of nog echt wel in de beginfase.

Spreker 2: Dat weet ik niet, ik overzie eigenlijk alleen de domeinen waar ik van ben dus watermanagement en leefomgeving en crisismanagement en daarbij zie je dus dezelfde patroon, dus watermanagement doen we al een tijdje. Ook op dezelfde organisaties, dus het datamodel dat is wel al een tijdje neergezet. Alleen wat nieuw is dat we steeds intensiever moeten samenwerken. Dus dat is ook weer probleemgestuurd. Kijk dat water dat dat loopt al een paar duizend jaar door Nederland en de organisaties die het water beheren, die zijn de volgens mij ook de oudste overheidsorganisaties ter wereld, dus volgens mij is het hoogheemraadschap Delfland, waar ik zelf onder val is volgens mij de oudste publieke organisatie ter wereld die volcontinu in bedrijf is geweest. Die jongens die zijn al een tijdje bezig met het water, maar tot nog toe heeft niemand ze gevraagd om nou heel veel data met elkaar uit te wisselen, olie waterschappen tussen Rijkswaterstaat en zolang ze dat niet wordt gevraagd maakt het hen ook niet uit dat ze geen overkoepelend datamodel hebben. Maar nu wordt het leven anders, we hebben drie jaar geleden grote droogte gehad. Dat heeft een paar honderd miljoen aan schade opgeleverd aan de agrarische sector alleen al. Dan wordt het een probleem en dan beginnen mensen te vragen: hoe kan het zijn dat we door een stuk van Nederland water te veel hebben en in een ander stuk van Nederland watertekort, en dat kost ons heel veel geld. Dan wordt het overdragen van dat water en overdragen van water data een probleem, en dan gaan we ons druk maken over het datamodel. Dus het datamodel is onderdeel van een communicatieoplossing. Maar als niemand met elkaar wil communiceren, is het datamodel, het is leuk voor in de boekenkast en intellectueel verantwoord en je kan ook nog afstuderen als je wil, maar voor de rest boeit het verder helemaal niemand. Zodra dus druk op die organisaties komt te staan om uit te wisselen dan wordt een informatiemodel opeens heel erg belangrijk. Datzelfde zie je nu bij de omgevingswet. De omgevingswet die wil integrale afweging maken over de omgeving. Dus als jij ergens iets wil doen, ik wil een café starten, dan heb je te maken met geluid én je hebt te maken met water, watergebruik, afvalwater, zwarte parkeerplaats, alcoholvergunning en brand vergunning. Dat hoort er allemaal bij, en dat moet dus nu met de omgevingswet integraal bekeken worden. Er voor waar het allemaal aparte wetten, werd dus ook aparte databases en dan moet opeens allemaal bij elkaar komen. Dus al die aparte databases, die geen informatiemodel hebben daarboven, dat is opeens een probleem geworden. Daar zie je dus nu ook bewegingen in ontstaan. Het informatiemodel is nooit een doel op zich, het is altijd onderdeel van een communicatie vraag en de data uitwisseling vraag. Aks er geen data uitwisseling vraag is, dan is het informatiemodel een vrij leuke intellectuele exercitie maar verwacht er niet te veel van.

Spreker 1: Okay juist ja. Ik vroeg mij ook af, zijn er standaarden of requirements? Of gewoon een standaard vanuit Rijkswaterstaat waaraan de datamodellen moeten voldoen?

Spreker 2: Nee, die heb ik nog niet gezien.

Spreker 1: Nee? Er is geen standaard van-

Spreker 2: Nee. Nee. Nee. Nee. Het is niet gezegd van: gij zult uw datamodellen ten alle tijde in uw weldoen, of u zult ten alle tijden afdoen aan de vierde normaalvorm of nee ik heb er niks van gezien. Het komt ook, omdat je ziet dus dat Rijkswaterstaat nu nog bezig is met zich te realiseren dat ze überhaupt iets met die data moeten. Dus het feit dat ze dan iets met die datamodellen moeten, dat begint bij de voorlopers in het denken begint dat nu wel duidelijker worden, maar wat je nu vraagt is nog de overtreffende trap daarvan. Als je dan die datamodellen wil beheren dan moeten die datamodellen zelf ook weer in een standaard staan. Aan die vraag is Rijkswaterstaat nog even niet.

Spreker 1: Niet toe? Okay.

Spreker 2: Nee je moet eerst een aantal van datamodellen boven water hebben gekregen, dat een aantal organisaties onderdelen zeggen: hé wij van deze organisatie gebruiken dit datamodel en degene die dat niet doet krijgt een vreselijke straf. Dus dan weet je ook, een datamodel is geland en dan kun je zeggen oke dan is het beheren van datamodel en de standaard van een datamodel dat wordt dan belangrijk. Maar dat is stap twee pas. Dat betekent wel dat mensen zoals jij en Margreet Kamp met andere data architecten, die moeten natuurlijk wel twee stappen vooruitlopen op de meute. Dus is het wel handig als zij daar over nadenken en daar dingen ook opschrijven van hoe ik zo'n datamodel standaard zou willen neerzetten. Daar moet je over nadenken, maar ik zou er nog niet veel over communiceren want het land nu niet.

Spreker 1: Iets voor later, als alles iets verder is.

Spreker 2: Voor iets voor later en misschien een voorloper die denkt van: hé, ik ben nog wel vrij ver met mijn datamodel van mijn stukje, laat ik het ook eens conform die standaard doen. Dat is ook een interessante oefening, want zodra de mensen zeggen: ja hé, we hebben nou tien datamodellen én eigenlijk kunnen we het niet goed beheren. Dan is er vast één iemand die het datamodel wel conform de standaard kan laten zien. Die kan laten zien: kijken die standaard die werkt ook daadwerkelijk want, ik heb hem gebruikt en het werkt nog steeds. Nou dan gaan we die standaard wel adapteren. Dus even, aan de ene kant zeg ik van ja die datamodellen en datamodel standaard, daar zijn we nog niet zo van. Van de andere kant is het zeker voor een afstudeerder heel nuttig om toch die vingeroefeningen te doen omdat je de zaak al klaar legt op het moment dat er wel vraag naar is, want als je pas begint met na te denken op het moment dat die vraag er is dan schuif je zo twee jaar op. Dan verlies je echt tijd.

Spreker 1: Er wordt toch wel al een klein beetje over nagedacht?

Spreker 2: Ja toch, en zeker voor wat wij als Rijkswaterstaat ik interessant vinden voor het afstuderen is we daar ook weer eventjes een inzicht krijgen in hoe de kennisinstituten, hbo en de universiteiten, hoe die naar dit soort vraagstukken kijken want die zijn in hun denken natuurlijk vier stappen verder dan Rijkswaterstaat. Wij willen wel die vier stappen vooruit, die willen we wel nog een keertje zien. Want anders dan komen wij als Rijkswaterstaat ook in onze tunnelvisie terecht en blijven we doen zoals we dat al dertig jaar doen. Na een tijdje zit je nou opeens in een heel verouderd denkbeeld en is de rest van de wereld je voorbijgegaan. Dan heb je de boot gemist, dat moeten we niet doen. Daarom zijn afstudeerders voor ons heel belangrijk om ook weer eventjes bijgepraat te worden over de laatste inzichten, in de theorie van hoe hoor je dit nou een beetje knap te doen. En we denken oké dit is een leuk inzicht uit de theorie en de praktische implementatie ervan die komt misschien pas over vijf tot tien jaar, maar dat is niet zo erg. Want je bent ook nog vrij jong. Over vijf tot tien jaar ben jij nog steeds aan het werk én kun je nog steeds gebruik maken van de kennis die je dan opdoet dus dat is geen probleem.

Spreker 1: Zeker waar. Okay, maar jullie hebben geen standaard. Hebben jullie dan wel al een soort van requirements waar de datamodellen aan moeten voldoen?

Spreker 2: Nee, ook niet.

Spreker 1: Ook niet? Dus het maakt niet uit hoe het er bijvoorbeeld uit ziet?

Spreker 2: Naar mijn weten niet nee. We hebben dus onze informatie in water, die heeft dan een datamodel neergezet in UML. Dat is wel een beetje een soort van defacto standaard, dat is iets waar we soort van ingerold zijn, maar dat is nooit afgesproken. En kijk, het is niet zo dat wij niks aan standaard doen dus op zich als je zegt van welke standaarden gebruiken wij nou om applicaties neer te zetten op technische infrastructuur? Die standaard die hebben wel. Welke producten gebruiken wij standaard om webservice die deployen, daar hebben we heel lijstje van. Van deze producten moet je gebruiken. Dat is een standaard en daar zit een architectuurraad op en die keurt het ook formeel goed. Dus we spreken wel degelijk standaarden af, maar niet op dit gebied.

Spreker 1: Het is nog te jong om eigenlijk zulke dingen te doen.

Spreker 2: Ja, maar dat is ook wel één van de dingen die je aan Margreet [Kamp] kan meegeven als onderdeel van je afstudeeropdracht. Dat is wel opvallend als je nou Rijkswaterstaat vergelijkt met wat er in de boeken staat van hoe je het zou moeten doen, dus dat je op dat gebied helemaal geen standaarden hebt. Terwijl, en dat is natuurlijk wel toegevoegde waarde voor het afstuderen, die er vanuit de wetenschap wel degelijk zijn.

Spreker 1: Ja toch wel, we hebben wel geleerd. Hier moet een datamodel aan voldoen, maar wel interessant dat het jullie gewoon niet zoveel uitmaakt qua hoe het eruit, om zo te zeggen.

Spreker 2: Nou goed of we doen het wel maar het is in ieder geval niet opgeschreven.

Spreker 1: Het is gewoon niet echt, de eisen?

Spreker 2: Nee, nee, nee, bij mijn weten is er geen sturing op datamodellen en ik kan je ook zo gauw geen- er is ook geen tool afgesproken waarin we hebben gezegd als u dan een datamodel heeft hoort die in die tool te staan. Hebben we ook niet, misschien is het chaos van datamodellen maar heb je de chaos bij elkaar in één tool kun je geval bekijken

Spreker 1: Want er worden nog wel steeds andere tools gebruikt behalve sparx ik had begrepen dat we ook iets van, iets ouds. Ik weet niet meer hoe het heet. Ik had het ergens genoteerd, wat jullie vijf jaar terug hadden gebruikt voordat jullie Sparx hadden?

Spreker 2: Troux.

Spreker 1: Ja dat inderdaad. Komt dat nog steeds voor?

Spreker 2: Volgens mij is dat nog steeds niet formeel uitgezet, dat klopt ja.

Spreker 1: Okay, dus er zijn nog steeds wel echt verschillende modelleer tools waarin het wordt gezet?

Spreker 2: Ja, hoewel Troux wel een tool is die echt nu bevroren behoort te zijn. Ik durf niet hardop te zeggen dat Troux niet meer actief wordt gebruikt, maar daar is wel degelijk afspraak over. Dat hoort niet. Er zitten een aantal administraties in Troux die we nog steeds nodig hebben dus die demonstraties zelf die worden bevroren maar het is niet de bedoeling dat je nu een Troux nog nieuwe dingen start.

Spreker 1: Alleen oude dingen die daar al in staan.

Spreker 2: Alleen die dan toevallig er nog in zaten, die nog niet gemigreerd zijn, die zitten er nog in. Ik durf niet honderd procent te garanderen dat er naast Sparx en Troux geen andere tools zijn waar datamodellen in zouden kunnen zitten. Maar ik denk dat Sparx nu wel de leidende tool is. Zeker binnen de CIV.

Spreker 1: Nou als we het toch over Sparx hebben, weet je toevallig waarom er ook echt gekozen is voor Sparx als tool. Ik had wel begrepen dat het eigenlijk niet als eerste functie had voor het modelleren van dingen maar voor andere dingen en dat is gewoon een goede bijkomst van sparx was

Spreker 2: We hebben Sparx gekocht- Wat wij wilden hebben is een tool voor de architecten en daarvoor hadden wij Troux een Troux vonden we te duur. Dus wilden we een nieuwe tool aanschaffen. Dat is een onderdeel binnen Rijkswaterstaat die door industrie automatisering bouwstenen moduleert. Niet dat er veel gunstige ervaringen zijn opgedaan met Sparx. Bovendien is Sparx één van de grotere marktleiders en voor het aanschaffen van tools is ons devies altijd: volg de marktleider want dat is de grootste kans dat hij niet halverwege failliet gaat.

Spreker 1: Zeker waar. In hoeverre is de data architectuur en data management- Ik weet niet eigenlijk moet ik dit eerst vragen. Hebben jullie daar eisen voor? Voor datamanagement en data architectuur of dat ook niet?

Spreker 2: Dat weet ik niet dat zou je aan Margreet Kamp en aan het datamanagementcentrum zelf moeten vragen want die gaan erover.

Spreker 1: oké nou dan vraag ik het aan hen. Even kijken. Iets anders wat mij was opgevallen. In de opdracht moest ik eigenlijk alles in UML modelleren. Maar dat vond ik wel bijzonder want UML, voor zover ik weet is eigenlijk bedoeld als vóór stap voordat je een database gaat opstarten eigenlijk. Want je kan natuurlijk, ja dan moet ik even goed zeggen, volgens mij functies in doen van dit hoort in deze functie. Dit wordt gedefinieerd in deze functie, maar eigenlijk bestaat de database al. Het is net een soort van andersom gebeuren. Waarom is er alsnog gekozen voor UML, want eigenlijk heb je alleen maar klassen, attributen nodig en dan de relaties maar dat kon eigenlijk denk ik ook in een ander model, in plaats van UML.

Spreker 2: Dat weet ik niet waarom de informatieanalisten voor UML hebben gekozen. Nee, dat weet ik niet. Ik heb het ook alleen maar te horen gekregen omdat de architect die hebben gezegd wij werken in Archimate en de informatieanalisten zeiden zo van wij willen ons datamodel graag maken in UML omdat dat dat dichter tegen het relationele database schema aan zit en mogelijkheden biedt om rechtstreeks de database schema te genereren.

Spreker 1: Ja dat is wel zo ja.

Spreker 2: Ik heb het niet nagerekend, ik vond de argumentatie wel plausibel Archimate is als architectuurtool nog behoorlijk conceptueel en abstract. Voor architecten is dat wel leuk, maar als je echt een technisch werkend database wil genereren is dat gewoon een te hoog abstractieniveau en heb je liever modelleer tool die daar wat dichter op de techniek zit. Waarom ze nou precies daar voor UML hebben gekozen en niet wat anders, dat weet ik niet.

Spreker 1: Okay, misschien dat margreet deze vraag kan beantwoorden.

Spreker 2: Ja, ik zou het even aan margreet vragen.

Spreker 1: Ik wilde ook nog even kijken naar DAMA DMBOK maar ik vroeg me af; in jouw optiek welke gebieden zijn dan handig om daar rekening mee te houden?

Spreker 2: Waar wil je naar kijken?

Spreker 1: Naar DAMA DMBOK dame het kent elf kennisgebieden en het gaat eigenlijk over data governance. Dus voor elk soort kennisgebied heeft het zeg maar één soort van; dit moet je volgens dit framework doen.

Spreker 2: Ja dat is het databeheer model wat ze hebben op laten stellen toch

Spreker 1: Ja het is inderdaad een soort van standaard voor data governance of eigenlijk databeheer

Spreker 2: Ja dat heb ik wel voorbij zien komen, maar dat heb ik niet echt precies bekeken wat de elf functies daarin precies zijn, omdat mijn aanname was dat IGA [Inwinning & Gegevens analyse] al die functies doen. Daar werk ik niet dus laat ik die problemen graag aan de andere directie even over.

Spreker 1: Weet je toevallig wie ik daarna kan vragen

Spreker 2: Dat is wel makkelijk van het hebben van een data architect. Dat zou ik ook wederom aan Margreet [Kamp] moeten vragen.

Spreker 1: oké nee, dan is de rest allemaal voor Margreet [Kamp]. Dat waren eigenlijk echt alle vragen. Super bedankt, het geeft mij ook weer meer inzicht en alles

10.3.3 Interview Rob Smits

Datum van interview: 13-04-2022

Geïnterviewde: Rob smits

Geïnterviewt door: Bo-Lan Schouten

Spreker 1 Bo-Lan Schouten

Spreker 2: Rob Smits

Spreker 1: Ja, ik heb eigenlijk gewoon een hele vragenlijst en die ga ik gewoon een beetje door. Zou je misschien eerst zelf kunnen voorstellen? Naam en functie.

Spreker 2: Ja prima. Ik ben Rob Smits en ik werk bij de Rijkswaterstaat. De dienst informatievoorziening. En, ik ben daar een strategisch adviseur en ik heb ook de rol van informatiemanager.

Spreker 1: Okay, top. Hoe gaat Rijkswaterstaat om met de data voordat er datamodellen waren?

Spreker 2: Nou ja, data is natuurlijk het allerbelangrijkste in de organisatie. Rijkswaterstaat bouwt weliswaar kunstwerk en legt wegen aan, maar daar gaat dus heel veel data vooraf. Om dat allemaal te kunnen realiseren. Dus data is heel belangrijk. Dat beseffen mensen vaak niet. Aan de andere kant, is het zo dat binnen de Rijkswaterstaat ja, iedereen heeft zelfs een eigen informatie bij elkaar. Alleen datamodellen en dergelijke worden belangrijk als je dus informatie verzamelt en vanuit verschillende hoeken van de organisatie voor het management. Want dan moet je dus informatie of data die het hetzelfde zijn bij elkaar pakken, maar als iedereen daar verschillend invulling aan geeft, verschillende definitie gebruikt, dan kun je ze niet zomaar bij elkaar pakken. Dan wordt het vrij arbeidsintensief om ze bij elkaar op te tellen, dan worden datamodellen belangrijk, dan wordt het belangrijk om definities te gaan maken. Van oké als we over een project praten, wat is dan een project? Als we over een begin van een project praten, wat is het begin van een project, wat is het eind van het project? En dat soort gegevens moet je dan definitie instellen voordat je dus ook dan informatie bij elkaar kan zetten en voor het hogere management beeld geven over alle projecten van Waterstaat. En dat je dan ook over hetzelfde praat.

Spreker 1: In de opdracht kwam naar voren dat er een definitieprobleem is en eigenlijk zijn de datamodellen juist bedoeld om dat definitieprobleem een beetje aan te pakken, toch?

Spreker 2: Ja.

Spreker 1: Maar hoe zijn die definitie problemen in het verleden opgelost. Dat je echt gewoon, ja zoals ja al zei, de startdatum. Hoe wordt dat uiteindelijk wel gewoon recht getrokken van: Dit is de startdatum?

Spreker 2: Nou ja, dan, dan krijg je dus. Heel veel werk eigenlijk, dus iemand heeft opgeschreven. De een gebruikt een andere startdatum dan de ander, en dan moet je dus of, je gaat er overheen en zegt van oké, er zitten verschillen in, maar dat accepteren we. Maar dan krijg je dus onkeurigheid in je gegevens. Of je hebt heel veel werk omdat je dus ja, je krijgt uit het systeem opgeleverd. Dit is de nieuwe data, maar dan moet je dus gewoon controleren of dat klopt met wat je werkelijk wil hebben en dan moeten we ook weer een andere. Als het ging even niet klopt, moet dat weer opgezocht worden. Wat het juiste gegeven is, dus je zag gewoon dat het tot nu toe. Dan is het vrij arbeidsintensief om de optelsom te maken.

Spreker 1: Maar waarom kwamen eigenlijk al die definitieproblemen, waarom niet vanaf begin af aan gekozen, dit is echt een definitie.

Spreker 2: Nou ja, je ziet dat. Zoiets groeit eigenlijk. En ja dan iedereen legt gewoon zijn gegevens zo goed mogelijk vast, maar iedereen doet dat. Ja, hoe hij en zij of het gewend is. Én als je dan, eigenlijk is er dan van bovenaf te weinig gezegd van zo moet het.

Spreker 1: Er waren niet echt eisen waaraan iemand moet voldoen om. Bijvoorbeeld om dingen vast te leggen?

Spreker 2: Nee, nee, nee, maar je ziet een beetje kenmerkend voor de RWS organisatie dat er veel meer vanaf de werkvloer naar boven komt dan dat van boven wordt opgedragen hoe de werkvloer Moet aandacht geven, want er is ook te weinig aandacht voor die data. Iedereen vindt dat vanzelfsprekend. Op het moment dat je dus echt uit systemen informatie wil gaan halen en dat wil gaan presenteren. Dan komt er eigenlijk veel meer nadruk te liggen op de definities van gegevens. Ik zeg altijd, de computers zijn hartstikke dom, dus die pakken gewoon op wat er staat. In het verleden werd dat door Mensen gedaan en Mensen zijn veel slimmer, want die zeggen van Oh, wacht, dat klopt niet.

Spreker 2: Die werkte dat direct even bij en dat doet de computer niet nodig of Je moet daar hele programma's voor gaan schrijven. Dus wat voorheen door mensenwerk, al een beetje vanzelfsprekend werd gecorrigeerd. Op het moment dat je kan digitaliseren, echt met computers gaat werken, wordt het veel kritischer om precies dezelfde gegevens te gebruiken, want hij telt gewoon bij elkaar op, terwijl twee verschillende dingen zijn.

Spreker 1: Maar hoe hanteer je dan nu? Zeg maar dat de datakwaliteit goed blijft. Omdat je zei dat het eerst met de hand wordt gedaan, Maar dat nu nog steeds met de hand gedaan, dan alsnog?

Spreker 2: Ja nu, je gaat nu ook gegevens presenteren op een dashboard. En die dashboards ja dan maak je, enerzijds een dashboard om je management informatie te presenteren. Als daar fout in zitten, dan valt het wel op dat er iets raars gepresenteerd wordt, dus dan krijg je al een vraag van, wat is hier aan de hand? Maar aan de andere kant maken we ook nog een keer een ander soort dashboard om te kijken of die gegevens wel consistent zijn, dus We gaan ook met onze interne dashboards die data aanmaakt. Die gaat ook nog kijken van, ja kloppen die gegevens wel met elkaar. Is het ook logisch? Dus ja, je gaat zelf een data controleren en als er iets gepresenteerd wordt wat raar is, dan zul je vragen krijgen.

Spreker 1: En hoe beheert of onderhoud Rijkswaterstaat eigenlijk de datamodellen?

Spreker 2: Ja, Dat is een goeie vraag, dus daar zijn we nu mee bezig. Dus hoe gaan we dat nou goed doen? Ja, Dat is de grote vraag. We gaan nu serieus aan de slag om die conceptuele datamodellen aan te maken en ja, dus dan komt ook de vraag van ja en hoe gaan we dat goed beheeren en wat komt er allemaal bij kijken? Dat zijn een beetje de vragen die nu voorliggen hier bij de organisatie.

Spreker 1: jullie hebben nog niet echt een soort van: Dit is hoe we doen.

Spreker 2: Nee.

Spreker 1: Dat komt ook een beetje omdat jullie nog niet echt datamodellen hadden. Want dat nog vrij nieuw is oké, toch?

Spreker 2: Ja.

Spreker 1: En wat gebeurt er momenteel aan het modelleren van de datamodellen die staan nu natuurlijk in Word documenten, maar- Of in hoeverre zijn jullie al klaar of nou ja klaar, met het moduleren van al die Word documenten? En het opstellen van die documenten.

Spreker 2: Nou ja, goed, We hebben dus wel gekeken naar van. Nou ja, We moeten ergens vastleggen, wat kun je het beste voor gebruiken? We hebben Sparx Enterprise, is een tool die we aangekocht hebben voor architectuur. En, nou ja, dan blijkt dus dat het tool nog meer kan, dus ook datamodellering. Dus dat ja, die hebben dus nu even direct ook erbij gepakt. Maar daarnaast zijn we ook met een programma bezig. Centrale toegang data, en nou is dit ook een onderdeel onder data governance. Dus naast mij is men ook aan het kijken van ja wat? Als we die data goed willen beheren, wat komt er allemaal bij kijken? De vraag die dan ook in de boven komt is, van ja in welke tooling heb je nodig om dat goed te kunnen doen. Dus we zijn nog druk bezig met allerlei programma's en met mensen in gesprek te zijn om over te kijken naar, ja data hoe en wat komt er allemaal bij kijken? Wat hebben we dan nodig?

Spreker 1: Dat is nog niet helemaal in kaart gebracht al.

Spreker 2: Nee, nee, nee. Ja precies Dat zijn we nu middenin om dat goed in beeld te krijgen. Dus dan ja, Dat is ook de vraag die bij jou ligt, van oké, hoe kunnen we dit nou het beste beheren, en ja, dat kunnen we straks gaan toepassen.

Spreker 1: Ik hoop dat het goed gaat komen. Is er vanuit Rijkswaterstaat een standaard waar datamodellen aan moeten voldoen? Of zijn er requirements of eisen of wensen ofzo waar ze aan moeten voldoen, afgezien van UML.

Spreker 2: Nee, We hebben wel een data architect, hè? Margreet Kamp? En ja aan haar stel ik de vraag, van welke staan daar dan hebben en ja, ik heb begrepen. dat er wat betreft data of dat met modellen te maken heeft, weet ik niet precies, maar er zijn heel wat modellen. Waar wij onder andere mee te maken hebben in de organisatie, is een datamodel- eigenlijk heet het informatiemodel: ruimtelijke ordening. We hebben natuurlijk ook de basisregistraties binnen het rijk. Basis registratie voor gebouwen, voor ja adressen en dergelijke. Dus daar zit ook een model onder, er zijn dus verschillende modellen. Margreet Kamp als data architect heeft ook voor de objecten van Rijkswaterstaat ook een basisvoorzieningen modellen opgebouwd. In de GEO-wereld worden ook allerlei data gebruikt. Dus daar zijn modellen voor geformuleerd. Je hebt ook nog een programma dat heet Erbinnen, dat is het bouw informatiemodel dus dat is ook al een model. Dus er zijn wel verschillende typen modellen die ook weer net andere doelen dienen. Dus het is ook zo de vraag van oké. We pakken die in elkaar en wat zijn de ja, hoe sluit dat op elkaar aan en waarom moeten we daarmee omgaan? Dus dat is wel een vraag die naar boven komt, nu we dat meer en meer aan op allerlei plekken dit soort modellen toegepast worden.

Spreker 1: Die modellen die er nu al zijn dat is een soort van de basis voor soortgelijke modellen, dus als we kijk naar bijvoorbeeld modellen over de gebouwen dat je zei, in principe zou datzelfde wel nog kunnen toepassen op een ja andere, stel je voor je gaat andere vestigingen ofzo doen, dan zou hetzelfde model daarvoor kunnen gebruikt worden, dus dat dan de standaard ervoor.

Spreker 2: Ja misschien. Ja, daar moeten we naar kijken, dus het Kadaster heeft een basisregistratie gebouwen. Ja, Dat is een basisregistratie, nationale basisregistratie, dus daar ga je van uit. Daar zitten

dus bepaalde dingen vastgelegd, dus daar ga je van uit. Je hebt ook een normen. Iso normen. Daar ga je ook de probeer zoveel mogelijk aan te sluiten. Dan heb je nog afspraken die binnen de branche worden gemaakt. De aanleg van de grond weg en waterbouw branche of noem maar op of je hebt de aankomende richtlijnen uit Europa. Dat zijn allemaal dingen waar je naar moet kijken. Rekening moet houden, dus Dat is een beetje waarvoor ik naar de data architecten kijk van ja, wat speelt er allemaal? Hoe zit dat in elkaar? Daar moeten we natuurlijk op aansluiten.

Spreker 1: Maar die norm, dat zou best wel ja nationale en internationale normen. Zijn er ook zeg maar binnen Rijkswaterstaat echt zo van: dit willen we per se in elk model erin hebben zitten.

Spreker 2: Ja dus als daar echt dingen erg belangrijk zijn en dan gaan we dat vastleggen in een kader. Hè, dus we hebben ook Rijkswaterstaat in kaders, en dan leggen we dat daarin vast.

Spreker 1: En, hebben jullie zeg maar wensen en eisen met betrekking tot de data architectuur aan datamanagement, dus hoe dat gedaan wordt? Zijn er eisen aan hoe de datastructuur eruitziet? En hoe het datamanagement gedaan wordt?

Spreker 2: Nou ja, Margreet Kamp heeft al iets opgeschreven over hoe we, ja hoe we met datastructuur omgaan, maar dat is nog beperkt eigenlijk. We willen wel in de volle breedte beter gesteld zijn, er moet nog wel wat meer gebeuren denk ik.

Spreker 1: oke maar een kleine opzet ervan als het ware?

Spreker 2: Nou ze heeft een stuk geschreven over data governance, ze heeft een stuk geschreven over doen- Maar we hebben een nieuwe CIO, hè? Dus een aantal onderdelen, die zijn al bij hem ingebracht, ten eerste dus breed in de organisatie vast laten leggen. Maar dat zit nog ja, de data strategie is eigenlijk nu twee jaar oud, denk ik. Daar vindt dus een uitwerking van plaats, een data agenda is dat eigenlijk en daar zitten allerlei stappen in om dit goed te vast te leggen, dus ja, daar gaan we vanuit. Ja.

Spreker 1: En als we kijken naar het plan dat er staat of vastgelegd staat én kijken naar de werkelijkheid, komt ook een beetje met elkaar overeen?

Spreker 2: Ja, dat weet ik nog wat informatie over binnengekeken. Wat ik wel weet van onze hele organisatie is dat het- Je ziet altijd heel verschillen tussen allerlei onderdelen. Dus ik verwacht dat we dus, ja er ook hier met data veel verschillen zullen zijn. Uit het verleden, weet ik ook wel dat ik. Ik heb wel eens met mijn district hoofden gesproken over het vastleggen over informatie over de wegverlichting. En de ene had het over wegverlichting en de ander had het over lantaarnpalen. En de derde had zelfs 5 verschillende typen van lantaarnpalen. Dus ja, ieder district keek er op zijn eigen manier naar en zei van: ja ik vind het wel goed hè? Er zit belijning op de weg, dat vind ik prima en de ander wist precies welke verf er was gebruikt voor de belijning, dus iedereen ging er op een andere manier mee om met de Informatie over zijn wegvak, zeg maar. Dat verwacht ik dus nu met de data dadelijk ook, dat iedereen dat op een eigen manier doet. De ene wil heel gedetailleerd, iets weten en de ander, zegt van: Nee, als ik het ongeveer weet, vind ik het prima en dan met de aannemer overleg ik wel wat er precies moet komen, als het zover is. Nou ja.

Spreker 1: In beide situaties zou het ook gewoon kunnen, of zouden dan toch wel meer mensen gewend zijn om meer in detail te treden of meer oppervlakkig te houden.

Spreker 2: Waarom het gaat, is van- wat ik al zei, de noodzaak om tot datamodellen te komen was omdat je dus managementinformatie wil genereren. En al die details dat is te veel voor het management. Je kunt ergens op een lager niveau veel details hebben en als je wat meer naar hogere niveaus gaat kun je gewoon details achterwege laten. Dat is een beetje hoe dat zit. En dan kun je dus op lager niveau kun meer details inderdaad hebben. Dus als een districtshoofd 5 verschillende lantaarnpalen wil onderkennen, prima. Maar op het niveau van zijn directeur weten we gewoon. Nou ja, dat is een weg met wegverlichting. Klaar. En dan op een nog hoger niveau heb je het gewoon over wegen en niet meer over de wat er allemaal bij zit, aan wegverlichting en noem maar op. Hoe hoger je komt in de managementlagen en de rapportage daarover, hoe minder details en uiteindelijk rapporteren we zelfs aan de Tweede Kamer. Daar staat gewoon in van er is 3000 km snelweg. Maarja, niet elke snelweg is hetzelfde natuurlijk en er zit ook weer heel veel details bij die wel bekend zijn op lagere niveaus, maar ook In de Tweede Kamer is dat allemaal niet meer van aan de orde.

Spreker 1: Je zie tot stand zijn gekomen, omdat het management zeg maar niet aan kon als het ware?

Spreker 2: We willen dus nu eigenlijk dashboard genereren. Wat is dashboard, en dashboard is het digitaal eigenlijk ophalen van informatie uit databases en dan komt hetzelfde punt wat ik net aangaf computers zijn hartstikke dom. Dus Je moet heel precies vertellen wat ze, waar ze op moeten halen, en wat ze op moeten halen om dan bij elkaar op te tellen en dan weer te presenteren. Dat moet met elkaar kloppen Als het niet met elkaar klopt, dan krijg je dus ja rare dingen. En wat ik al zei, vroeger deden mensen dat met de hand en ja, die zagen direct als het niet klopt en dan zetten ze met de hand het juist cijfer erin. Als je dat automatiseert, digitale automatiseert en dan moet je daar, dan moet het wel met elkaar kloppen en dan komt dat veel nauwkeuriger. Dan heb je die datamodellen nodig, want dan moet je dus echt zorgen dat het goed in de database staat omdat je dan met programma als SaS en dergelijke die informatie ophaalt uit die databases en daaruit een dashboard genereert.

Spreker 1: Waarom is dan eigenlijk juist nu gekozen om datamodellen te visualiseren? Waarom is dat niet eerder gebeurd?

Spreker 2: Ja de datamodellen zijn nu pas klaar.

Spreker 1: En, waarom is er niet eerder begonnen om een datamodellen te maken?

Spreker 2: Nou ja, Omdat ook die trend van die dashboards nu pas, afgelopen jaren op gang komt dus wat ik al zei, we gaan nu echt het bewust ophalen, direct geautomatiseerd uit systemen. Dat is echt recent.

Spreker 1: Dat hebben jullie nog niet echt heel veel eerder gedaan, dus er was ook geen behoefte aan.

Spreker 2: Niet geautomatiseerd hè? Dus het werden wel een soort dashboard gebouwd, maar dat was allemaal handwerk.

Spreker 1: Dat is veel werk.

Spreker 2: Ja, weliswaar met een Excel spreadsheet of zo. Maar dat deed iemand. Ja, die liep door dat Excel spreadsheet je heen en ja, dan werden dingen ook eventjes met de hand bijgewerkt. Dat hele arbeidsintensieve verdwijnt nu en maar je hebt ook nu dat dus ja, de actualiteit hè, dus je maakt er dan één keer per jaar of twee keer per jaar, maak je een rapportage. Dan werd bijgewerkt en dan was men daar even mee bezig. En nu heb je dus ja, als er nu een vergadering is, dan kan iemand van het

management kan zo even doorklikken naar de website om te kijken naar het dashboard. Die eigenlijk een heel actueel beeld geeft op dat moment. Omdat de database zeer ja-, als je dus voortdurend die gegevens bijhoudt, hè, dan krijg je heel actuele beeld. En ja zijn de cijfers veel nauwkeuriger, dus tot nu toe werden dus wel beslissingen genomen, maar ja, dan eigenlijk op allerlei ja niet zo betrouwbare data.

Spreker 1: Ja oke. Wordt er in de gehele organisatie datamodellen vastgelegd, of verschilt echt per afdeling?

Spreker 2: Ja, ja hier in dit gebouw is het meer begonnen, zeg maar. Maar dat is ook wel vanzelfsprekend omdat dus hier meer mensen zitten die van data en ICT verstand hebben, dus. Die begrijpen wat noodzakelijk is. Ook omdat hier de Mensen zitten die dashboards bouwen, dus die vragen er ook om. En bij op andere plekken? Zie je nu ook dat men ja- Dat reduceert van ja wacht even, het is wel handig om een datamodel hebben want tot nu toe klopt het allemaal niet. Ze zeggen nu van: hé wacht, even er klopt iets niet. Hoe kan dat nou? En dan krijg je dus het antwoord vanuit deze kant, de CIV van ja, dat komt omdat je afspraken niet vastgelegd zijn, want dat kan je doen door een datamodel te maken. We hebben nu ook een RWS-project database, hè? Die wordt al jarenlang gebruikt, centraal bij de diensten die het project draaien. GPO is dat en PPO. En ja nu ook is er een dashboard gemaakt voor het vervanging en renovatieprogramma. En wat zien ze nu, dat het dashboard soms rare cijfers laat zien. Hoe komt dat nou? Omdat er dus geen goed datamodel ligt onder die RWS-projecten database. Ze constateren dus hè, dankzij dashboard: he er gebeuren rare dingen. En dan blijkt dus dat de gegevens niet eenduidig zijn.

Spreker 1: Dat is dan nooit eerder opgevallen dan behalve bij de dashboards?

Spreker 2: Ja dus niet eerder opgevallen Omdat er geen dashboard was. Het programma vervanging en renovatie waar die is toegepast, is ook nieuw. Tot nu toe, ja op die manier is dat hele programma een nieuwe aanpak die onlangs is opgestart met het dashboard erbij. Dus ja, daardoor komt eigenlijk naar boven toe.

Spreker 1: Oké bedankt, Ik had gisteren ook met Margreet Kamp een interview én zij zei ook nog dat sommige afdelingen geen datamodellen vastleggen en ik vroeg me af Waarom hun geen datamodellen vastleggen? Heel veel doen het ook wel, maar een paar doen het ook niet. Zij wisten hier geen antwoorden, want ja, Misschien dat jij hier wel een antwoord op had.

Spreker 2: Nou, ja, het gaat meer en meer worden dus een paar jaar geleden deed niemand het. We zijn nu- Ja, We hebben nu noodzaak gezien. We zijn ermee begonnen.

Spreker 1: En in andere afdeling speelt dat ook?

Spreker 2: Andere afdeling zullen Misschien dadelijk ook nog wel komen. Ik denk dat het-. Kijk in de nieuwe strategie van Rijkswaterstaat staat ook dat we een data gedreven organisatie willen worden dus het denken in data, dat is vrij nieuw nog binnen de organisatie. Dan komen dit soort dingen allemaal naar boven toe van ja. Als je een data gedreven organisatie bent, wat betekent dat nou? Dat betekent onder andere dat je dus ook. Meer aandacht geeft aan de data, maar ook meer aandacht geeft aan de kwaliteit van de data. En dan is het weer van, ja, maar wat spreken we dan over die data? En dat gebeurt dat in een datamodel. En nog niet iedereen ziet het belang van een datamodel. Dat besef zal de komende jaren meer en meer komen dan ja dan zullen ze dus ook een conceptueel datamodel doen. Dus Ik heb ook een memo geschreven voor het directieteam hier, dat er dus ja, echt behoefte is aan een groepje mensen. Die weet, hoe hoort het eigenlijk te zijn zo'n datamodel? Wat

zijn daar de eisen aan? En als ze datamodellen maken, dat ze ook kunnen zorgen dat de datamodellen goed op elkaar aansluiten. En dat we allemaal op dezelfde manier, ook met die datamodellen omgaan. Nou goed wat ik al zei van nee, we moeten ook het beheer nog inrichten en er moet ook een groepje mensen komen die dat gaan doen.

Spreker 1: Dat heb je nog niet?

Spreker 2: Nee, nee, nee. We hebben eigen alleen maar ja-, we hebben dus Victor [van der Kloet] die dat dan deed. En daarvoor was ook nog iemand die het even heeft gedaan. Maar ja, er zijn nog niet echt functies vastgelegd in de organisatie die daar ja echt voor aangewezen zijn. Ja Margreet Kamp doet het ja. Die heeft eigenlijk een andere taak, maar ze deed het er een beetje erbij. Daar zullen we gewoon meer aandacht aan moeten geven. Ja, dat vraagt wel ook voor het personeel wat. Daar heb je dan mensen voor nodig. Die weten hoe dat moet. Die dat ook helemaal gaan doen en dan andere mensen helpen om dat datamodel te maken.

Spreker 1: Het valt mij ook wel op dat het allemaal in UML moest, maar volgens mij is UML een beetje bedoeld voor als je nog geen database hebt. Dan ga je een kijken: Oké, dus de data die ik in mijn organisatie heb, Dit heeft met elkaar zo een relatie en vervolgens ga je eigenlijk op basis daarvan een database maken, want je kan in UML ook functies/operations toevoegen en dan heb je dus een klasse en dan zeg je, deze klasse heeft de volgende functies, dus die maakt bijvoorbeeld iets aan of die wijzigen iet of iets dergelijks. En dan wordt dat dus uiteindelijk een database, maar jullie hebben natuurlijk alle databases. Waarom is alsnog gekozen voor UML?

Spreker 2: Want Dat is het advies van ik kreeg. Want ja, Waarom Misschien wel? Wel, ja, dat maakt het mij ook allemaal niks uit. Het signaal wat ik kreeg, was van ik vroeg eigenlijk van ja, die datamodellen staan nu in een in een Word document. Het lastige is dat je dus de relaties daardoor moeilijker goed in zicht kreeg. En ook als je dus op één plek iets veranderd, dat je ook weet dat op een andere plek ook de verandering wordt doorgevoerd. En, ja, soms heb je dat je dus een dingetje in het geheel veranderd. En dan ja, dan moet je eigenlijk weer een nieuw document maken. Terwijl Als je dat in zo'n database achtige-. Of ja In het systeem hebt zitten met alle informatie, en je verandert iets, dan is het hele verhaal is nog steeds compleet, dan hoeft je niet een nieuw systeem te maken. Dus voor het bijhouden en zeker voor allerlei kleinere mutaties. Ja had ik wel een wens dat er een systeem komt om dat te gaan doen. En toen kreeg ik het advies van, ja, dat kunnen we prima in Sparx Enterprise doen. Sparx Enterprise heeft UML, dus we zouden het heel mooi in UML kunnen doen. Als dat handig is; zeg ik prima.

Spreker 1: Daar is meer vanuit de tooling gekeken naar, wat is handig?

Spreker 2: Nou ook vanuit het advies die ik kreeg vanuit de architectuur. Omdat het prima past om het in UML te doen en dat het ook aansluit met wat er al is. Als je het in conceptueel datamodel hebt, dan kunnen we het zo doorzetten naar het gebruik van dashboards of het inrichten van databases.

Spreker 1: Okay zo dus. Uhh- ja dat waren alle vragen. De andere-, ik had er nog twee over, maar die zijn al redelijk goed beantwoord dus, ja dat was het, super bedankt voor de tijd en de antwoorden.

Spreker 2: Ja graag gedaan.

10.4 Bijlage D: Workshop Sparx en UML

Training EA en UML dag 1



Figuur 20: Bewijs van deelname workshop

10.5 Bijlage E: Functionaliteiten per editie Sparx Enterprise Architect

In onderstaand overzicht zijn alle functionaliteiten per editie te zien (Sparx Systems, 2022b).

	Ultimate	Unified	Corporate	Professional
Ada 2012 Code Engineering	✓	✓		
Advanced UML 2.5 Modeling	✓	✓	✓	✓
Allocated Work	✓	✓	✓	✓
ArcGIS	✓	✓	✓	✓
ArchiMate 2.0	✓	✓	✓	✓
Archimate 3 Support	✓	✓	✓	✓
Audit Model Changes	✓	✓	✓	
Auto Create Diagram Image Map	✓	✓	✓	✓
Automation API	✓	✓	✓	✓
AUTOSAR	✓	✓	✓	✓
Azure icon set NEW	✓	✓	✓	
AWS icon set v15	✓	✓	✓	
BABOK & BIZBOK v14	✓	✓	✓	
Baseline Diff/Merge	✓	✓	✓	
BPEL Generation from BPMN diagrams	✓	✓		
BPEL 2.0 Generation	✓	✓		
BPMN 2.0	✓	✓	✓	✓
BPMN Simulation	✓	✓		
BPMN XML	✓	✓	✓	✓
BPSim Modeling & Simulation Built-in v14	✓	✓		
BMM Profile v15	✓	✓	✓	
Business Process Modeling	✓	✓	✓	✓
Business Rules Composer	✓	✓		
Case Management (CMMN) v14	✓	✓	✓	
Cephes Math Library NEW	✓	✓	✓	
Chart External Models	✓	✓	✓	
Charts and Dashboards	✓	✓	✓	✓
Cloud Connectivity with Firebird (feap) repositories v14	✓	✓	✓	✓
Cloud Connectivity	✓	✓	✓	

Code Engineering					✓
					✓
Construct View v15	✓	✓	✓	✓	✓
Context Browser v14	✓	✓	✓	✓	✓
Custom Drawing Style v15	✓	✓	✓	✓	
Custom Progress Bars	✓	✓	✓	✓	✓
Custom Reports v15	✓	✓	✓	✓	
Custom Tables v15	✓	✓	✓	✓	
Customized Model Patterns	✓	✓	✓	✓	✓
Dashboard Views NEW	✓	✓	✓	✓	✓
Data Miner v15	✓	✓	✓	✓	
Database Engineering	✓	✓	✓	✓	✓
Database Engineering DB Builder	✓	✓	✓	✓	
DBMS Repository**	✓	✓	✓	✓	
Debugging of Jscript, VBScript and JavaScript	✓	✓	✓	✓	
Debug & Visualize Applications	✓	✓	✓	✓	✓
Debug using JDWP protocol	✓	✓	✓	✓	✓
Decision Model and Notation (DMN) - Modeling v14	✓	✓	✓	✓	
Decision Model and Notation (DMN) - Simulation v14	✓	✓	✓	✓	
Default Document Handler	✓	✓	✓	✓	✓
Diagram Discussions v14	✓	✓	✓	✓	✓
Diagram Filters - Filtering on Tagged Values	✓	✓	✓	✓	✓
Diagram Inline Specification View v14	✓	✓	✓	✓	✓
Diagram Matrix View v15	✓	✓	✓	✓	✓
Diagram Theme/Visual Style	✓	✓	✓	✓	✓
Diagram Legends Auto Coloring	✓	✓	✓	✓	✓
DMN Decision Service v15	✓	✓	✓	✓	
Docked Property Windows (all major modeling components) v14	✓	✓	✓	✓	✓
Document Fragments in Reporting	✓	✓	✓	✓	✓
Dynamic Charts NEW	✓	✓	✓	✓	
Dynamic Documents	✓	✓	✓	✓	✓
Ecore	✓	✓	✓	✓	✓
Element Discussions	✓	✓	✓	✓	
Element List (Tabular Editing)	✓	✓	✓	✓	✓
Entity Relationship Diagram (ERD)	✓	✓	✓	✓	

Executable Code Generation from behavioral Models	✓	✓		
Executable State Machines	✓	✓		
Execution Analyzer Menu	✓	✓	✓	✓
Export Model Reference Data for Calendar events, Team Review, Gap Matrix Profiles, and Patterns	✓	✓	✓	✓
External Data Source Integration # v14	✓	✓	✓	
Feature Matrix v15	✓	✓	✓	✓
Firebird Repositories	✓	✓	✓	✓
Floating Edition Available	✓	✓	✓	✓
Gantt Charts	✓	✓	✓	✓
Gap Analysis	✓	✓	✓	
GML 3.3 and 3.2.1	✓	✓	✓	✓
Google Cloud Platform icon set v15	✓	✓	✓	
Hand Drawn Mode	✓	✓	✓	✓
Heat Maps	✓	✓	✓	✓
Hybrid Scripting	✓	✓	✓	✓
Info View	✓	✓	✓	✓
Integrated Project management using Kanban	✓	✓	✓	✓
Interaction Flow Modeling Language (IFML) v14	✓	✓	✓	✓
Kanban Diagrams	✓	✓	✓	✓
Lazy Load	✓	✓	✓	
Learning Center	✓	✓	✓	✓
Linux & macOS installation via WINE	✓	✓	✓	✓
Linux satellite service	✓	✓		
MARTE Profile v15	✓	✓		
Math Support built into Script Engines	✓	✓		
MATLAB Support (via Solvers) NEW	✓	✓	✓	
MDG Technologies (Create)*	✓	✓	✓	✓
MDG for AML	✓	✓	✓	✓
Metadata/Repository Advanced Search	✓	✓	✓	✓
Metamodel Views v14	✓	✓	✓	✓
Mind Mapping	✓	✓	✓	✓
Model Based Add-ins v15	✓	✓	✓	
Model Chat	✓	✓	✓	
Model Driven Architecture (MDA)	✓	✓	✓	✓

Model Pattern Library	v14	✓	✓	✓
Model Repository (EAP & FEAP Project Files)		✓	✓	✓
Mono debugger	v14	✓	✓	✓
Native XML export	v15	✓	✓	✓
Navigation Bar		✓	✓	✓
New Report Templates		✓	✓	✓
NIEM		✓	✓	✓
NIEM 4	v14	✓	✓	✓
Octave Support (via Solvers)	NEW	✓	✓	✓
Office 2016 Visual Style		✓	✓	✓
OMG XMI		✓	✓	✓
Package Menu		✓	✓	✓
Perspectives	v14	✓	✓	✓
Personal Daily Journal	v14	✓	✓	✓
Portals Window		✓	✓	✓
Profile Authoring Helpers		✓	✓	✓
Program Memory Profiler	v14	✓	✓	✓
Program Stack Profiler	v14	✓	✓	✓
Project Browser Favorites	v15	✓	✓	✓
Project Management, Maintenance and Testing Windows		✓	✓	✓
Project Gantt View		✓	✓	✓
PDF Report Generation		✓	✓	✓
Personal Tasks View		✓	✓	✓
Profile/Metamodel Extensibility		✓	✓	✓
Project Calendar		✓	✓	✓
Project Tasks		✓	✓	✓
Publish to Joomla!		✓	✓	✓
Real-Time, HDL Code Engineering and Profiles		✓	✓	✓
Relationship (Traceability) Matrix		✓	✓	✓
Relationship Matrix Textual Overlays		✓	✓	✓
Replicate .EAP Projects		✓	✓	✓
Report Generation: HTML and Rich-Text		✓	✓	✓
Report Customization: WYSIWYG Template Editor (for Rich-Text Reports)		✓	✓	✓

Reverse Engineer User Defined Languages	✓	✓	✓	✓
Reusable Asset Service (Read)	✓	✓	✓	✓
Reusable Asset Service (Write)	✓	✓	✓	
Roadmap Diagrams	✓	✓	✓	✓
Roadmap Patterns & Profile v14	✓	✓	✓	✓
Ribbons Interface	✓	✓	✓	✓
Risk, Task and Image Objects	✓	✓	✓	✓
RTF Reporting API	✓	✓	✓	✓
State Machines Code Generation	✓	✓		
Save Diagrams as PDF	✓	✓	✓	✓
Schema Composer	✓	✓	✓	
Scripting with JScript, VBScript and Javascript	✓	✓	✓	
Security (Role-based)	✓	✓	✓	
Security Aware Perspectives v15	✓	✓	✓	
Semantic Information Modeling for Federation SIMF v14	✓	✓	✓	
Shape Scripts/Customization	✓	✓	✓	✓
Shared Models	✓	✓	✓	✓
Simscape Integration NEW	✓	✓		
Simulation of Object Instances	✓	✓	✓	
Simulink Integration NEW	✓	✓		
Solver Classes NEW	✓	✓	✓	
SOMF 2.1	✓	✓	✓	✓
Specification Manager	✓	✓	✓	✓
Statechart Editor	✓	✓	✓	✓
Stateflow Integration NEW	✓	✓		
SysML 1.4, 1.3, 1.2, 1.1	✓	✓		
SysML 1.5 - Modeling v14	✓	✓	✓	
SysML 1.5 - Simulation with OpenModelica v14	✓	✓		
SysML Parametric Model Simulation	✓	✓		
SysML Simulation with OpenModelica	✓	✓		
SysPhS Support NEW	✓	✓	✓	

Task Allocations				
Team Library	v14			✓
Test Management	✓	✓	✓	✓
Testpoints	✓	✓	✓	✓
Testpoints for VB.NET	✓	✓	✓	✓
Time Aware Modeling	✓	✓	✓	✓
Total User Interface Themeing	✓	✓	✓	✓
TOGAF Gap Analysis	✓	✓	✓	
UML Model Simulation (Activity, Sequence, State) ***	✓	✓	✓	✓
User Interface Simulation	✓	✓	✓	
Value Delivery Modeling Language (VDM) v15	✓	✓		
Version Control Integration	✓	✓	✓	✓
Version Control TFS 2010, 2012 Support	✓	✓	✓	✓
Virtualized Ends	✓	✓	✓	✓
Visual Studio / Eclipse Links	✓	✓	✓	✓
Win 32 User Interface Designs	✓	✓	✓	✓
Wireframes	✓	✓	✓	
Working Sets	✓	✓		
WSDL Engineering	✓	✓	✓	✓
XMI Import and Export (2.1, 1.2, 1.1, 1.0)	✓	✓	✓	✓
XML and XSD Editing and Validation	✓	✓	✓	✓
XMI Merge using Baseline Merge sets	✓	✓	✓	
XML Schema (XSD) Engineering	✓	✓	✓	✓
XSLT Debugger	✓	✓	✓	✓
WAN Optimizer	✓	✓	✓	

RPSim Modeling & Simulation

DDS	✓	✓		
DOORS Link	✓	✓		
Erlinse Integration	✓	✓		
Erlinse Link	✓	✓	✓	✓
Microsoft Office	✓			
SysMI	✓	✓	✓	
TOGAF	✓	✓		
Visual Studio Integration	✓	✓		



10.6 Bijlage F: Afstudeervoorstel

DATAMODELLERE N VOOR ICT- ASSETS EN IT- PROJECTEN STAGEVOORSTEL

Naam: Bo-Lan Schouten

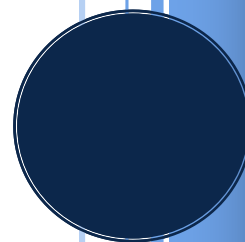
Studentennummer: 1115829

Opleiding: Informatica

Specialisatie: Business data management

Afstudeerorganisatie: Rijkswaterstaat

Datum: 15 december 2021



INHOUDSOPGAVE

VOORWOORD.....	2
SAMENVATTING	3
FIGURENLIJST	6
TABELLENLIJST.....	6
BEGRIPPENLIJST.....	7
1. INLEIDING.....	8
2. BESCHRIJVING STAGEORGANISATIE.....	9
2.1 Stagebedrijf.....	9
2.1.1 Organogram.....	9
2.2 Stagebegeleiding.....	10
2.3 Stakeholders	10
3. PROBLEEMSTELLING.....	12
4. DE CONTEXT	13
4.1 Het proces.....	13
4.2 De data architectuur	15
4.2.1 FAIR principes.....	16
5. HET ONDERZOEK	17
6. METHODOLOGIE	18
7. ONDERZOEKSRESULTATEN – DEELVRAGEN.....	20
7.1 Is EA geschikt voor het modelleren van de gewenste UML-modellen?.....	20
7.2 Wat zijn de requirements waar de modellen aan moeten voldoen?	21
7.3 In hoeverre is door het bedrijf dit probleem omtrent modelleren van conceptuele datamodellen al behandeld?.....	23
7.4 Met welke ‘regels’ moet er rekening worden gehouden bij het modelleren van een UML Klassendiagram?	26
7.5 Wat zegt DAMA DMBOK over het onderhouden/beheren van data(modellen) ?	29
7.6 Wat zegt De Cloud Data Management Capabilities Framework over het onderhouden/beheren van data(modellen) (Hier naar kijken)!	34
8. BEROEPSPRODUCT	36
8.1 De datamodellen.....	36
8.2 Het advies	39

8.2.1	<i>Datamodellen ontwerp</i>	39
8.2.2	<i>Data management framework</i>	41
9.	REFLECTIE	44
9.1	Competenties	44
9.1.1	<i>A-Competenties</i>	44
9.1.2	<i>B-Competenties</i>	46
9.2	Planning	48
9.3	Uitkomsten	49
9.4	Begeleiding	49
10.	BIJLAGE	51
10.1	Bijlage A: Literatuurlijst	51
10.1.1	<i>Afbeeldingen</i>	51
10.1.2	<i>Websites</i>	51
10.1.3	<i>Boeken</i>	52
10.1.4	<i>PDF-documenten</i>	53
10.2	Bijlage B: Het bedrijfsobjectmodel RWS (BOMR)	54
10.2.1	<i>Het huidige bedrijfsobjectmodel RWS (BOMR)</i>	54
10.2.2	<i>Het bedrijfsobjectmodel RWS (BOMR) 2009</i>	54
10.3	Bijlage C: Interviews	55
10.3.1	<i>Interview/vragenlijst Victor van der Kloet</i>	55
10.3.2	<i>Interview Peter Bernhard</i>	56
10.3.3	<i>Interview Rob Smits</i>	66
10.4	Bijlage D: Workshop Sparx en UML	73
10.5	Bijlage E: Functionaliteiten per editie Sparx Enterprise Architect	74
10.6	Bijlage F: Afstudeervoorstel	81
1.	ORGANISATORISCHE ASPECTEN	85
1.1	Gegevens stagebedrijf	85
1.2	Gegevens bedrijfsbegeleider	85
1.3	Gegevens Studieloopbaan begeleider	85
1.4	Gegevens stagiaire	85
2.	CONTEXT/ACHTERGROND VAN HET AFSTUDEERBEDRIJF	86
3.	HET PROBLEEM	87
4.	BEDRIJFSDOELSTELLING	87
5.	CONCRETE OPDRACHTOMSCHRIJVING	87
6.	DELIVERABLES (PRODUCTEN)	88
7.	PLANNING	89

8. VERDEDIGING A-COMPETENTIES	90
8.1 De A-Competenties	90
8.2 Definiëren van de A-Competenties	90
8.3 Verdediging A-Competenties	90
9. B-COMPETENTIES	91
9.1 De gekozen B-Competenties	91
9.2 Definiëren van de gekozen B-Competenties	91
9.3 Verdediging van de B-Competenties	92
9.2.1 Verdediging opdracht in context	92
9.2.2 Verdediging adviseren data oplossing	93
9.2.3 Verdediging analyseren data architectuur	93
BIJLAGE: BRONVERMELDING	95
10.7 Bijlage G: Adviesrapport Dataframework en datamodellen	96

1. ORGANISATORISCHE ASPECTEN

1.1 Gegevens stagebedrijf

Stagebedrijf: Rijkswaterstaat Centrale Informatievoorziening (CIV)

Adres: Derde Werelddreef 1, 2622 HA Delft.

Website: <https://www.rijkswaterstaat.nl/>

1.2 Gegevens bedrijfsbegeleider

Naam: Rob Smits

Telefoon: 06 5142 1498

Mail: rob.smits@rws.nl

1.3 Gegevens Studieloopbaan begeleider

Naam: Chaimae Lamaakchaoui

Telefoon: 06 2311 7888

Mail: Lamaakchaoui.c@hsleiden.nl

1.4 Gegevens stagiaire

Naam: Bo-Lan Schouten

Telefoon: 06 3109 2403

Mail: s1115829@student.hsleiden.nl

2. CONTEXT/ACHTERGROND VAN HET AFSTUDEERBEDRIJF

Rijkswaterstaat houdt zich voornamelijk bezig met infrastructuur en waterstaat. Het is verantwoordelijk voor beheren van grote wateren maar ook voor het onderhouden van dammen, dijken, stuwen en stormvloedkeren. Daarnaast is Rijkswaterstaat ook verantwoordelijk voor het onderhoud en de kwaliteit van de rijkswegen. Het is dan ook de uitvoeringsorganisatie van het ministerie van infrastructuur en Waterstaat

Rijkswaterstaat kent zeven landelijke en zeven regionale organisatieonderdelen. Tijdens mijn stage ben ik onderdeel van de dienst Centrale Informatie voorziening (CIV). Zij houden zich voornamelijk bezig met de ontwikkeling en beschikbaarheid van alle informatie. Zo zorgen ze enerzijds voor industriële automatisering bij bijvoorbeeld bruggen en tunnels, anderzijds houden ze zich ook bezig met kantoorautomatisering. Ze leveren programma's, producten en diensten voor alle onderdelen binnen Rijkswaterstaat.

Het organogram van Rijkswaterstaat ziet er zo uit:



(Rijkswaterstaat, z.d.)

3. HET PROBLEEM

Rijkswaterstaat kent vele processen en het verandert ook steeds. Hierdoor zijn er ook meerdere conceptuele datamodellen binnen Rijkswaterstaat. Momenteel heeft Rijkswaterstaat conceptuele datamodellen ten behoeve van hun IT-projecten en de ICT-assets. Deze zijn in opdracht van de informatiemanager van het proces Informatievoorziening tot stand gekomen. Echter staan deze conceptuele datamodellen in een word-bestand, waardoor dit zeer onoverzichtelijk is en moeilijk te onderhouden.

Daarnaast staan gegevens bijvoorbeeld ook lokaal bij iemand in een excel bestand. Hierdoor treedt het probleem op dat onder andere objecten een andere betekenis hebben voor verschillende personen. Er mist nu vooral uniformiteit waardoor het allemaal onduidelijk is. Het datamodel moet als het ware nog genormaliseerd worden zodat begrippen duidelijker worden en mensen makkelijk gegevens kunnen toevoegen.

4. BEDRIJFSDOELSTELLING

De conceptuele datamodellen staan nu in een Word document, maar ze willen graag dat dit wordt opgenomen in een Sparx Enterprise Architect. Het doel hiervan is dat de data (onder andere monitorings-, sturings- en verantwoordingsinformatie) dan beter te onderhouden is en ook meer overzichtelijk en inzichtelijk is. Ze verwachten tevens ook dat door de conceptuele datamodellen vast te leggen in Sparx Enterprise Architect de kwaliteit van de conceptuele datamodellen verbeterd wordt, doordat de (vele) relationele verbindingen tussen de data beter inzichtelijk is.

Daarnaast is ook het doel dat de er eenduidigheid wordt gecreëerd binnen de data. Omdat het nu ook in word staat en soms nieuwe processen bij komen of oude veranderen is dit niet het geval op het moment. Doordat er onder andere definities worden vastgelegd, wordt het makkelijker om dashboards voor sturing en verantwoording in te richten.

5. CONCRETE OPDRACHTOMSCHRIJVING

De opdracht is dus als volgt: Momenteel zijn er verschillende conceptuele datamodellen, deze moeten met elkaar afgestemd worden zodat er een duidelijk en overzichtelijk datamodel komt. De verschillende objecten moeten een definitie krijgen, dus wanneer wordt iets beschouwd als dat object. Hierdoor kunnen misverstanden ontweken worden. Het is hierbij dus voornamelijk de bedoeling dat dit de kwaliteit van de data bevordert en het meer inzichtelijk wordt.

Er zal een advies gegeven worden voor een nieuw datamodel. Vervolgens moet er een vertaalslag van de huidige conceptuele datamodellen die nu in Word staan naar UML-format maken en deze uiteindelijk vastleggen in Sparx Enterprise Architect. Gaande weg wordt er nog voorstellen gedaan voor besluitvormingen over aanpassingen in het bestaande conceptuele datamodel op basis van bevindingen of veranderingen bij de CIV-dienstverlenende processen. Ook wanneer er bijvoorbeeld mogelijke (ontwerp)problemen optreden kan er een advies komen hoe iets veranderd kan worden om dit probleem op te lossen.

Uiteindelijk zal er ook een soort van overdrachtsdocument/handleiding komen, zodat wanneer mijn stage bij Rijkswaterstaat ten einde komt, andere mensen dit nog wel kunnen gebruiken.

De stage wordt als voldaan gezien wanneer:

- Het UML model in Sparx Enterprise Architect staat
- Het UML model in Sparx Enterprise Architect te beheren en onderhouden is
- Alle competenties die door school zijn opgesteld, aangetoond zijn tijdens de stage

6. DELIVERABLES (PRODUCTEN)

Aan het einde van de afstudeerstage komen de volgende producten tot stand:

- UML model dat geïmplementeerd is in Sparx Enterprise Architect
- Handleiding/toelichting op dit model, met betrekking tot onderhoud
- Voorstellen voor besluitvorming over aanpassingen in de bestaande modellen

7. PLANNING

De stageperiode is van februari 2022 tot juli 2022. Hierbij zijn de (school)vakanties in acht genomen. Deze planning is een indicatie voor wanneer er een begin wordt gemaakt aan een bepaald onderdeel. Er kan afgeweken worden van deze planning. De planning is als volgt:

Bo-lan Schouten		project start	dinsdag 1 februari 2022																													
		project end	vrijdag 1 juli 2022		Weeknummer																											
TASK	PROGRESS	START	END	DAYS	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27					
Deliverables school																																
Afstudeerplan	0%	7-feb-22	11-feb-22	5																												
Evaluatie afstudeerplan	0%	14-feb-22	18-feb-22	5																												
Bedrijfsbezoek	0%	7-feb-22	11-feb-22	5																												
Opleveren afstudeerverslag voor intervisie	0%	11-apr-22	15-apr-22	5																												
tussen evaluatie	0%	18-apr-22	22-apr-22	5																												
Pre-final versie afstudeerverslag	0%	23-mei-22	27-mei-22	5																												
oplevering afstudeerverslag	0%	13-jun-22	17-jun-22	5																												
Introductie																																
Introductie bedrijf	0%	1-feb-22	4-feb-22	4																												
Kennismaking collega's en stakeholders	0%	1-feb-22	11-feb-22	11																												
Afstudeeropdracht																																
Na gaan hoe processen binnen RWS lopen	0%	7-feb-22	18-feb-22	12																												
Data stromen binnen RWS analyseren	0%	14-feb-22	18-feb-22	5																												
De huidige data architectuur analyseren	0%	21-feb-22	25-feb-22	5																												
De definities vastleggen van de verschillende (data) objecten	0%	28-feb-22	11-mrt-22	12																												
Verder analyseren van de huidige conceptuele datamodellen	0%	14-mrt-22	18-mrt-22	5																												
Nadenk over een mogelijk beter model	0%	21-mrt-22	1-apr-22	12																												
Mogelijk advies geven over een beter model	0%	4-apr-22	15-apr-22	12																												
De conceptuele datamodellen afstemmen	0%	4-apr-22	8-apr-22	5																												
Vertaal slag maken naar UML	0%	18-apr-22	6-mei-22	19																												
De datamodellen in Sparx zetten	0%	9-mei-22	27-mei-22	19																												
Advies schrijven over de zaken	0%	21-mrt-22	10-jun-22	82																												
Overdrachtsdocument maken	0%	30-mei-22	10-jun-22	12																												
Afronden	0%	5-jun-22	1-jul-22	27																												
Overig																																
Advies geven over mogelijke problemen	0%	14-mrt-22	1-apr-22	19																												

8. VERDEDIGING A-COMPETENTIES

8.1 De A-Competenties

Tijdens de afstudeerstage wordt er aan de volgende A-Competenties gewerkt:

- Onderzoek
- Leren leren
- Professioneel werken
- Innovatie

8.2 Definiëren van de A-Competenties

Toelichting – Onderzoek:

Onderzoek is een vertaling van de A-Competentie “A3 (Wetenschappelijke) Toepassing [onderzoek]” en “A7 Methodisch en Reflectief denken en handelen”, waarin we de nadruk willen leggen dat elk afstuderen een onderzoekstraject bevat waarin alle aspecten van het onderzoekproces terug komen. Waar dit proces ingebracht wordt en de omvang ervan kan per opdracht en specialisatie anders zijn (Hogeschool Leiden, 2021).

Toelichting – Leren leren:

‘Leren leren’ is is een vertaling van de A-competenties “A1 brede professionalisering” en “A4 transfer en brede inzetbaarheid”, waarin we de nadruk willen leggen opdat elke student in staat moet zijn zichzelf te vernieuwen wanneer nodig. Dit is een essentiële competentie in het ICT vakgebied dat met constante veranderingen te maken heeft. Als basis van ‘leren leren’ geldt het positieve zelfbeeld wat een student in staat stelt zelfredzaam te zijn en motiveert om nieuwe dingen op te pakken. Elke student moet sturing kunnen geven aan het zelfbeeld, moet zichzelf kunnen redden doordat hij de juiste tools kan toepassen en kan zichzelf motiveren om tot resultaten te komen (Hogeschool Leiden, 2021).

Toelichting – Professioneel werken:

Professioneel werken is een vertaling van de A-competenties "A6 probleemgericht werken", "A7 methodisch en reflectief denken en handelen" en "A8 Sociaal-communicatieve bekwaamheid" waarin de aanpak van het afstuderen centraal staat. De aanpak is een rode lijn door het hele afstuderen, wat start bij het probleem en eindigt bij de reflectie achteraf (Hogeschool Leiden, 2021).

Toelichting – Innovatie:

Innovatie, is boven de theorie staan, buiten de gestandaardiseerde reglementen en aanpak kunnen denken. Buiten de kaders dus. Mag inhoudelijk of op proces. Moet leiden tot iets nieuws, zoals een maker het betaamt. Innovatie is een vertaling van de A-competenties: “A4 Transfer en brede inzetbaarheid”, “A5 Creativiteit en complexiteit in handelen” en “A6 Probleemgericht werken” (Hogeschool Leiden, 2021).

8.3 Verdediging A-Competenties

Verdediging – Onderzoek:

Tijdens de stage zal er onderzocht worden naar hoe de modellen op de beste manier met elkaar

verbonden kunnen worden zodat dit inzichtelijker en overzichtelijker wordt. Er zal hierbij gekeken moeten worden naar verschillende manieren waarop dit kan gebeuren en onderzocht moeten worden wat de beste aanpak is. Deze aanpak zal dan uiteindelijk toepast worden in de praktijk.

Verdediging – Leren leren:

Rijkswaterstaat wil graag hun datamodel in Sparx Enterprise Architect komt te staan. Momenteel heb ik daar nog niet veel ervaring mee. Gedurende mijn stage zal ik dus ook leren om te gaan met Sparx Enterprise Architect. Daarnaast zal er nog dieper in worden gegaan op datamodelleren/data architecturen, en zal ik waarschijnlijk nog meer dingen leren naast de theorie uit de lessen.

Verdediging – Professioneel werken:

Tijdens de stage wordt er ook nauw samengewerkt met de informatieanalist van de afdeling Platformen. Er zullen overleggen plaatsvinden en daarnaast zal er ook gewoon meegedraaid worden in het proces. Er worden eventueel ook overige overleggen bijgewoond, etc. Daarnaast zullen er ook overleggen zijn met de stagebegeleider van het bedrijf, waardoor er genoeg ruimte is voor feedback. Dit zorgt er ook voor dat alles op schema blijft lopen.

Verdediging – Innovatie:

Momenteel zijn er al datamodellen, maar deze staan in Microsoft Word en zijn niet overzichtelijk. Een van de doelen van de stage is om de huidige datamodellen op zo'n manier te modelleren dat dit wel het geval is. Tijdens dit proces kunnen er verschillende problemen ontstaan/optreden. Hierbij zal er dan waarschijnlijk een idee bedacht moeten worden die dit oplost.

9. B-COMPETENTIES

9.1 De gekozen B-Competenties

Tijdens de afstudeerstage moeten drie B-competenties op niveau 3 aangetoond worden. Ik heb gekozen voor de volgende drie B-competenties:

- Opdracht in context plaatsen (Organisatieprocessen analyseren)
- Adviseren data oplossing (Organisatieprocessen adviseren)
- Analyseren data architectuur (Infrastructuur analyseren)

9.2 Definiëren van de gekozen B-Competenties

Toelichting – Opdracht in context plaatsen (Organisatieprocessen analyseren):

- De student past methoden en technieken uit de organisatiekunde toe om een organisatie te beschrijven. Deze analyse is er op gericht om context te geven aan de opdracht die de student uitvoert.
- De student beschrijft de huidige data architectuur.
- De student past methoden en technieken van data management toe om te omschrijven op welke huidige manier de organisatie met data omgaat.
- De student formuleert de eisen en wensen waar een data oplossing aan moet voldoen. Dit kunnen functional en non-functional requirements zijn als er een nieuwe of aangepaste data architectuur moet komen; of onderzoeksvragen als er een data analyse wordt uitgevoerd (Barthel, 2020).

Toelichting – Adviseren data oplossing (Organisatieprocessen adviseren):

- De student formuleert een advies.

- De student plaatst het advies in de context van een analyse van de organisatie zelf en de omgeving van de organisatie.
- De student onderbouwt het advies.
- De student past de relevante methoden en technieken van organisatie-en verandkunde, data modelleren, data-analyse, data management en data architectuur toe om tot een advies te komen (Barthel, 2020).

Toelichting – Analyseren data architectuur(Infrastructuur analyseren):

- De student beschrijft de huidige data architectuur.
- De student maakt technische datamodellen van de relevante bestaande data storages.
- De student kiest de functionele gebieden van data management die relevant zijn voor de organisatie en de opdracht.
- De student analyseert de huidige situatie voor ieder gekozen functioneel gebied van data management.
- De student onderzoekt in welke mate data architectuur en data management in lijn zijn met de eisen en wensen van de organisatie (Barthel, 2020).

9.3 Verdediging van de B-Competenties

9.2.1 Verdediging opdracht in context

Punt	Verdediging
De student past methoden en technieken uit de organisatiekunde toe om een organisatie te beschrijven. Deze analyse is er op gericht om context te geven aan de opdracht die de student uitvoert.	Om uiteindelijk een nieuwe model te kunnen maken zal eerste gekeken moeten worden naar de organisatie zelf. Zo kan er dan nauwkeurig worden nagegaan hoe de datastromen en zo definitie gegeven worden aan de data
De student beschrijft de huidige data architectuur.	De huidige modellen worden beschreven. Hieruit voort komt een nieuw, duidelijk en overzichtelijk model. Het huidige model zal wel de basis zijn voor het nieuwe model, daarom dat er heel goed gekeken moet worden naar het huidige model
De student past methoden en technieken van data management toe om te omschrijven op welke huidige manier de organisatie met data omgaat.	Een nieuwe model kan alleen tot stand komen wanneer er eerst wordt gekeken naar de data zelf. Wat betekend de data, en hoe gaat de organisatie met de data om. Daarnaast moeten de processen ook bekeken worden zodat dit afgestemd kan worden op elkaar. Dit zal gedaan worden middels methoden technieken en uiteindelijk ook vastgelegd worden
De student formuleert de eisen en wensen waar een data oplossing aan moet voldoen. Dit kunnen functional en non-functional requirements zijn als er een nieuwe of aangepaste data architectuur moet komen; of	Om een goed model op te leveren dat voldoet aan de verwachtingen van de organisatie zal er in samenspraak met stakeholders vastgesteld moeten worden waar zo'n model precies aan moet voldoen. De opdracht zal

onderzoeksvragen als er een data analyse wordt uitgevoerd.	dan vervolgens met het oogpunt op de requirements uitgevoerd worden
--	---

9.2.2 Verdediging adviseren data oplossing

Punt	Verdediging
De student formuleert een advies.	Tijdens de stage zullen er voorstellen gedaan worden met betrekking tot besluitvorming over aanpassingen in de bestaande modellen. Dit is dus een advies hoe iets anders gedaan kan worden, daarnaast zal er ook advies gegeven worden over hoe uiteindelijk een datamodel er uit kan/gaat zien
De student plaatst het advies in de context van een analyse van de organisatie zelf en de omgeving van de organisatie.	Dit advies zal natuurlijk op basis van de wensen en eisen van de organisatie gedaan worden. De voorstellen voor besluitvorming over aanpassingen zullen op basis van bevindingen of veranderingen bij de CIV-dienstverlenende processen gedaan worden.
De student onderbouwt het advies	Tijdens zo'n voorstel moeten er natuurlijk redenen genoemd worden waarom dit eventueel nodig is. Daarom dat het advies onderbouwd moet worden, anders is het niet duidelijk waarom dit advies opgevolgd moet worden.
De student past de relevante methoden en technieken van organisatie-en veranderkunde, data modelleren, data-analyse, data management en data architectuur toe om tot een advies te komen	Er wordt eerst gekeken naar de huidige stand van zaken, op basis hiervan komt een advies tot stand voor een nieuw model. Hierbij wordt er ook gekeken naar de technieken van data architectuur.

9.2.3 Verdediging analyseren data architectuur

Punt	Verdediging
De student beschrijft de huidige data architectuur	Om uiteindelijk een nieuw model tot stand te laten komen zal eerst de huidige modellen beschreven moeten worden. Uit de huidige modellen vloeit namelijk een nieuwe model voort
De student maakt technische datamodellen van de relevante bestaande data storages.	Tijdens de stage worden er dus nieuwe datamodellen opgesteld obv het huidige bedrijfsproces. Hier zullen dus ook data storages in meegenomen worden.
De student kiest de functionele gebieden van data management die relevant zijn voor de organisatie en de opdracht.	Bij het maken van een nieuw datamodel wordt er waarschijnlijk gekeken naar de volgende functionele gebieden van data management: Data Architecture, Data Modeling and Design en Data Quality. Mogelijk dat hier later meer gebieden bijkomen

De student analyseert de huidige situatie voor ieder gekozen functioneel gebied van data management.	Voorafgaand wordt dus eerst de huidige situatie geanalyseerd. Daarom zal er dus ook gekeken worden naar de functionele gebieden van data management, om zo de huidige modellen beter te begrijpen.
De student onderzoekt in welke mate data architectuur en data management in lijn zijn met de eisen en wensen van de organisatie	Vanuit de organisatie zijn er wensen/eisen waar een model aan moet voldoen. Eerst moet er natuurlijk gekeken worden naar de huidige modellen en nagegaan worden in hoeverre deze al naar de wensen van de organisatie is, vanuit daar kunnen verdere stappen ondernomen worden

BIJLAGE: BRONVERMELDING

Barthel, M. (2020, april). *B-competenties BDaM*. Hogeschool Leiden.

<https://hogeschoolleiden.sharepoint.com/sites/og2122.iwlab-afstuderen/Moduledocumentatie/Forms/AllItems.aspx?id=%2Fsites%2Fog2122%2Eiwlab%2Dafstuderen%2FModuledocumentatie%2FRubrics%2FToelichting%20B%2Dcompetenties%20BDaM%2Epdf&parent=%2Fsites%2Fog2122%2Eiwlab%2Dafstuderen%2FModuledocumentatie%2FRubrics>

Hogeschool Leiden. (2021, 20 augustus). *Informatica Rubrics A-competenties* [Presentatieslides].

SharePoint. <https://hogeschoolleiden.sharepoint.com/sites/og2122.iwlab-afstuderen/Moduledocumentatie/Forms/AllItems.aspx?id=%2Fsites%2Fog2122%2Eiwlab%2Dafstuderen%2FModuledocumentatie%2FRubrics%2FBijlage%20N%20%2D%20Rubrics%20A%2Dcompetenties%2Epdf&parent=%2Fsites%2Fog2122%2Eiwlab%2Dafstuderen%2FModuledocumentatie%2FRubrics>

Rijkswaterstaat. (z.d.). *Organisatiestructuur* [Organogram]. Rijkswaterstaat.

<https://www.rijkswaterstaat.nl/over-ons/onze-organisatie/organisatiestructuur>

10.7 Bijlage G: Adviesrapport Dataframework en datamodellen

Datum	25 april 2022
Versie	1
Status	DEFINITIEF

Colofon

Uitgegeven door

Auteur Bo-Lan Schouten
E-mail Bo-lan.schouten@rws.nl

Datum 14 juni 2022
Versie 1
Status DEFINITIEF

Versiebeheer

0.1	Concept	
1.0	Definitief	14-06-2022

Inhoud

	Inleiding	98
1	HET ONDERZOEK	99
2	ALTERNATIEVEN ONTWERP DATAMODELLEN	100
2.1	Het gekozen ontwerp	100
2.1.1	<i>Voordelen van deze manier:</i>	100
2.1.2	<i>Nadelen van deze manier:</i>	100
2.2	Het alternatief op het gekozen ontwerp	100
2.2.1	<i>Voordeel</i>	101
2.2.2	<i>Nadeel</i>	101
2.3	Conclusie/aanbeveling	101
3	ALTERNATIEVEN DATA MANAGEMENT FRAMEWORK	102
3.1	Requirements	102
3.2	DAMA DMBOK	102
3.2.1	<i>Waarom interessant voor Rijkswaterstaat</i>	102
3.2.2	<i>Wat houdt het in</i>	102
3.2.3	<i>Voordelen</i>	106
3.2.4	<i>Nadelen</i>	106
3.3	Data Management Capabilities	106
3.3.1	<i>Waarom interessant voor Rijkswaterstaat</i>	106
3.3.2	<i>Wat houdt het in</i>	106
3.3.3	<i>Voordelen</i>	107
3.3.4	<i>Nadelen</i>	107
3.4	Conclusie	107

Inleiding

In de periode februari 2022 – juni 2022 zijn er data ten aanzien van IV-projecten gemodelleerd conform UML en deze zijn visueel weergegeven in Sparx Enterprise Architect (EA). De reden voor het modeleren van deze modellen was zodat definitieproblemen verholpen werden en hierdoor betere en accuratere dashboards tot stand komen. Hierbij kwam echter ook de vraag hoe deze datamodellen vervolgens onderhouden moeten worden. In dit document wordt er keken naar twee ontwerpen voor klassendiagrammen en twee data management frameworks. De voor- en nadelen van elke optie wordt gedefinieerd en uiteindelijk komt hier een aanbeveling/advies tot stand.

1

Het onderzoek

De vraag was hoe de data(modellen) het beste onderhouden/beheerd kunnen worden. Het antwoord erop was echter niet zo makkelijk. Er was weinig literatuur te vinden op gebied van onderhoud datamodellen. Om uiteindelijk toch tot een goed advies te komen heb ik daarom naar twee dingen gekeken:

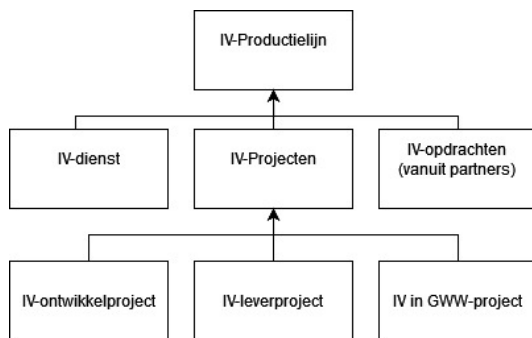
3. Tijdens het modelleren van de data: hoe kan dit vervolgens het makkelijkst onderhouden worden. Hierbij is er voornamelijk gekeken naar/vanuit de datamodellen.
4. Het probleem achter de datamodellen was als volgt: definities zijn onduidelijk → ingevoerde data heeft daardoor mogelijk niet dezelfde betekenis → gegevens op het dashboards zijn incorrect. Datamodellen het definitie probleem oplossing, echter is het nog wel de vraag of de ingevoerde data voldoet aan de beschrijving (de definitie) van een klasse. Is dit niet het geval dan heeft dit nog steeds gevolgen voor de dashboards. Om deze reden is er ook gekeken naar data management en met name data governance. Data governance zou er namelijk voor kunnen zorgen dat dit probleem opgelost wordt. Hiervoor is gekeken naar twee frameworks.

Om uiteindelijk tot een compleet mogelijk advies te komen is er dus gekeken naar de datamodellen en naar twee datamanagement frameworks. Er is gekeken naar het DAMA DMBOK en het CDMC™ (Cloud Data Management Capabilities) framework. Waarom er is gekozen voor deze twee frameworks en wat ze inhouden is in het volgende hoofdstuk te lezen.

2 Alternatieven ontwerp datamodellen

2.1 Het gekozen ontwerp

Bij het ontwerpen van de datamodellen is er voor gekozen om het in verschillende niveaus op te delen. Hierdoor is er eerst naar de bovenste laag gekeken dus de bedrijfsvoering. Hierin is te zien waar de IV-productielijn uit bestaat. In een nieuw diagram is er gekeken naar IV-projecten zelf. Hierbij is er gekeken naar wat er in het algemeen bij een IV-project komt kijken. Denk aan bijvoorbeeld budget, ongeacht welk type project hoort er altijd wel een budget bij. Hier zijn ook de 3 verschillende type projecten weergegeven, maar deze zijn verder niet uitgewerkt. Dit wordt gedaan op het volgende niveau. Hierbij is elk project afzonderlijk uitgewerkt. Klassen komen hierdoor dus wel in meerdere diagrammen voor, maar omdat er een kopie van een klasse gemaakt wordt is het dus niet het geval dat je dezelfde klasse op verschillende plekken moet bijwerken zodat deze ook in de andere diagrammen hetzelfde zijn. Bij het bewerken van een klasse, zullen alle kopieën ook automatisch bijgewerkt worden. Dus je past bijvoorbeeld de klasse IV-Projecten aan in het diagram IV-productielijn. Deze klasse staat ook vermeld in de diagram IV-projecten, maar omdat deze klasse in de diagram IV-project een kopie is hoeft deze dus niet apart aangepast te worden. Hieronder is een schematische weergave van hoe de datamodellen met elkaar zijn verbonden.



2.1.1

Voordelen van deze manier:

- Wanneer er een nieuw type project bijkomt dan kan dit in een apart diagram gemaakt worden waarin de project fases uitgewerkt worden. De hoofdklasse moet dan alleen toegevoegd worden aan het datamodel: IV-projecten. Het bijwerken van het model is hierdoor erg eenvoudig.
- Bij het implementeren van bijvoorbeeld een nieuw type project zal dit geen invloed hebben op de rest.
- Het is overzichtelijk, er staan niet 30 klassen in één model.

2.1.2

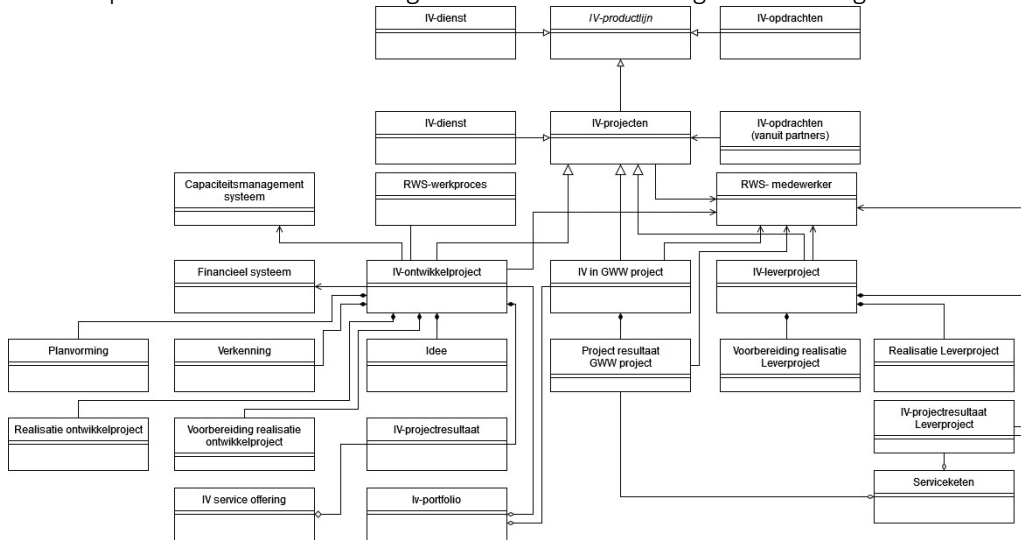
Nadelen van deze manier:

- Uiteindelijk kunnen er veel verschillende diagrammen ontstaan
- De mappenstructuur van hoe de diagrammen zijn opgeslagen is cruciaal om een logisch en hiërarchisch beeld te krijgen.
- Vanuit de bovenste twee lagen is er niet in één oogopslag te zien waar een project uit bestaat. Men zal hier eerst naar het juiste diagram moeten gaan

2.2

Het alternatief op het gekozen ontwerp

Echter was er ook een alternatief. Eventueel kon alles ook in één diagram komen te staan in plaats van alles op te delen in meerder diagrammen. Dat zou er ongeveer als volgt uitkomen te zien:



2.2.1 Voordeel

- In één diagram te zien waar een project bijvoorbeeld uit bestaat en hoe het is verbonden met andere projecten, zonder andere diagrammen erbij te hoeven halen.
- Klassen komen ook maar één keer voor in het hele diagram. Hier is bijvoorbeeld zo te zien dat twee van de drie projecten een relatie hebben met serviceketen.

2.2.2 Nadeel

- Veel Klassen waardoor het onoverzichtelijk wordt. Het duurt even voordat men alle entiteiten in zich heeft opgenomen.
- Heel veel relaties die vaak door elkaar heen lopen.
- Het toevoegen van een nieuw type project kan lastig worden omdat er ruimte gemaakt moet worden voor de klassen en de relaties.
- Het is even zoeken voordat je ziet welke klassen bij een project horen.

2.3 Conclusie/aanbeveling

Uiteindelijk is het gekozen ontwerp het meest praktische. Door dat het is opgedeeld in verschillende diagrammen is het kleiner en overzichtelijker. Het is ook makkelijker om nieuwe projecttypes/klassen aan bestaande projecttypes eraan toe te voegen omdat dit niet in een al volle diagram erbij gezet moet worden. Er kan namelijk gewoon een aparte diagram gemaakt worden voor een nieuw project of klassen aan een bestaand diagram toegevoegd worden. Als het aan een bestaand diagram toegevoegd moet worden dan gaat dit makkelijker in het huidige ontwerp dan wanneer alles in één diagram staat. Ook als er gekeken wordt naar UML best practices (Bellekens, 2018), dan komt het huidige ontwerp beter overeen met de punten die hierin genoemd zijn.

3 Alternatieven Data management framework

3.1 Requirements

Als er gekeken wordt naar requirements met betrekking tot dataheer dan zijn er 3 requirements waaraan het moet voldoen, deze staan vermeld in een kader dat betrekking heeft op databeheer (Kamp, 2021).

De requirements zijn als volgt. ,

- De governance rollen zijn ingevuld (bij voorkeur op basis van DROM-R)
- Er is een erkende eigenaar voor de data aangewezen o.b.v. zwaartepunt principe zie ook Bedrijfsobjectenmodel Rijkswaterstaat (BOMR)

3.2 DAMA DMBOK

3.2.1 *Waarom interessant voor Rijkswaterstaat*

Er is naar dit framework gekeken omdat het een van de bekendste data management frameworks is. Daarnaast kwam ook de vraag vanuit de organisatie om hiernaar te kijken en is ook bij de documentatie van de opzet van de conceptuele datamodellen aangegeven dat dit handig is om naar te kijken. Om deze redenen is er dus gekeken naar dit framework.

3.2.2 *Wat houdt het in*

DAMA DMBOK framework wordt gezien als een van de bekendste frameworks op gebied van datamanagement. Dit boek gaat over 11 kennisgebieden waarbij het draait om databeheer (SynTouch, 2020) (Bluefrog B.V., 2021) (Dama International et al., 2017). Het kent de volgende 11 kennisgebieden: Data governance, Data architecture, Data modeling & Design, Data Storage & Operations, Data security, Data Integration & Interoperability, Documents & Content, Reference & Master Data, Data Warehousing & Business Intelligence, Metadata, Data Quality.

Om het hele framework uit te werken kost te veel tijd en niet alles heeft betrekking op de scope, daarom dat er gekeken is naar 2 kennisgebieden: Data governance en Data architecture. Na gesprekken met de stakeholders en in het achterhoofd houdend dat de opdracht voornamelijk gaat over conceptuele datamodellen en onderhoud van deze modellen is er tot de conclusie gekomen dat deze twee functie gebieden het meest belangrijk zijn en binnen de scope vallen.

Data governance: Elk bedrijf heeft wel data, en maakt aan de hand deze data ook beslissing. Hierbij is alleen de vraag of zij ook een goede data governance hebben. Organisaties die dit namelijk wel hebben kunnen de waarde van hen data vergroten.

Binnen DAMA DMBOK wordt data governance als volgt gedefinieerd: Het is het uitoefenen van gezag en controle (planning, monitoring en handhaving) over het beheer van data assets. Dit is de basis en stuurt alle andere data managementfuncties aan. Het doel van data governance is dan ook om ervoor te zorgen dat data, in overeenstemming met beleid en best practices, op juiste manier worden beheerd. Het focust op hoe besluiten over data gemaakt worden en hoe mensen en processen verwacht worden te gedragen met betrekking tot data. De focus van een data governance programma is afhankelijk de behoeftes van de organisatie, maar zal meestal wel betrekking hebben op de volgende onderwerpen:

- **Strategy (strategie):** Het definiëren van, communiceren met en aansturen van de uitvoering van Data Strategy en Data Governance
- **Policy (beleid):** Het opstellen en handhaven van beleid met betrekking tot data en Metadata management, access, usage, security en quality
- **Standards and quality (standaarden en kwaliteit):** Het instellen en handhaven van normen voor data kwaliteit en data architectuur.
- **Oversight (toezicht):** Het bieden van hands-on observaties, audit en correcties op belangrijke gebieden van kwaliteit, beleid en data management (data stewardship).

- **Compliance (naleving):** Ervoor zorgen dat de organisatie kan voldoen aan de data gerelateerde wettelijke nalegingsvereisten
- **Issue management (probleem management):** Het identificeren, definiëren, escaleren en oplossen van problemen met betrekking tot data security, data access, data quality, naleving van regelgeving, data ownership, beleid, normen, terminologie of data governance procedures
- **Data management projects (data management projecten):** het steunen van pogingen op data management te bevorderen
- **Data asset valuation (waardering van data assets):** Standaarden en processen instellen om de business waarde van data te definiëren

Het is daarom belangrijk dat middels data governance policies en procedures ontwikkeld worden. Daarnaast is het ook belangrijk, dat het zich bezighoudt met organisatieveranderingsmanagement zodat dit ervoor zorgt dat de organisatie actief geïnformeerd wordt over de voordelen van verbeterde data governance en zo ook het gedrag heeft om data succesvol te beheren als een asset.

Om data dus succesvol te beheren als een bedrijfsasset is, is het ook belangrijk dat het bedrijf zelf de waarde van en datamanagement leert. Zelfs met de beste datastrategie, data governance en datamanagement plannen zal het niet slagen tenzij de organisatie de nodige veranderingen accepteert en beheert. De volgende principes kunnen helpen bij het leggen van een sterke basis voor data governance.

- Succesvolle data governance start met visionair en betrokken leiderschap. Datamanagement activiteiten worden bestuurd door een data strategie dat weer wordt aangestuurd door een enterprise business strategie.
- Data Governance is een business programma, daarom moet het zowel IT besluiten gerelateerd aan data beheren als de business interactie met data.
- In alle kennisgebieden van datamanagement is data governance een gedeelde verantwoordelijkheid tussen business stewards en technische datamanagement professionals.
- Data governance vindt door de hele organisatie plaats, op elk niveau.
- Omdat activiteiten op het gebied van data governance samenwerking tussen functionele gebieden vereisen, moet het data governance programma moet er een framework vastgesteld worden dat verantwoordelijkheden en interacties definieert.

Nu dat er een beeld is van wat er binnen de organisatie nodig om een goede data governance te hebben is de vraag nog, waarom. De twee hoofdredenen: risico's verminderen en processen te verbeteren.

Data governance kent ook verschillende activiteiten, hieronder worden ze uitgelegd:

Definiëren van data governance voor de organisatie: Het is hierbij de bedoeling dat het duidelijk wordt waar data governance over gaat, dus wat wordt er beheert, wie wordt beheerd en wie het beheert. Dit moet allemaal duidelijk zijn om het succesvol te maken. Dit kan gedaan worden door de gehele organisatie in plaats van naar bepaalde gebieden. Definiëren van de scope van data governance houdt meestal in dat het bedrijf zelf ook gedefinieerd wordt.

Uitvoeren van een beoordeling van gereedheid: Hierbij wordt de huidige status van de informatiemanagement capaciteiten beschreven. Volwassenheid en effectiviteit zijn cruciaal voor het plannen van een data governance programma. Dit houdt meestal in: begrijpen wat de organisatie doet met data; meet de huidige datamanagementmogelijkheden en capaciteiten. Daarnaast is het ook belangrijk om de mate van verandering binnen een organisatie te meten om er zo achter te komen waar weerstand kan ontstaan. Een beoordeling over samenwerking tussen management en datagebruikers is ook essentieel binnen deze activiteit. Er moet samengewerkt kunnen worden omdat dit meerdere domeinen betreft binnen een organisatie. Als er binnen een organisatie niet goed samen gewerkt kan worden kan dit een obstakel zijn voor stewardship. Als laatste zal er ook een

Business alignment assessment gedaan moeten worden. Hierbij wordt er gekeken naar hoe in lijn de organisatie is met hen business strategie, met betrekking tot data gebruik.

Uitvoeren van discovery en business alignment: Het data governance -programma moet bijdragen aan de organisatie door specifieke voordelen te identificeren en te realiseren. Discovery activiteiten zullen de effectiviteit van het bestaande beleid en richtlijnen identificeren en beoordelen. Discovery kan ook kansen voor data governance identificeren om zo de bruikbaarheid en inhoud van data te verbeteren. Daarnaast is ook datakwaliteit in onderdeel van discovery. Beoordelen van datakwaliteit zorgt ervoor dat er meer inzicht komt in de bestaande problemen en obstakels, maar ook zal het inzicht geven op wat de impact is van een slecht datakwaliteit en de risico dat men dan loopt.

Ontwikkelen data governance strategie: Een data governance-strategie definieert de scope en aanpak van data governance. De data governance moet uitgebreid gedefinieerd worden en in relatie gebracht worden tot de algemene bedrijfsstrategie, de datamanagementstrategie en de IT-strategie.

Ontwikkelen van goals, principes en policies: De ontwikkeling van doelen, principes en beleid afgeleid van de Data Governance-strategie zal leidend zijn voor de organisatie in de gewenste toekomstige staat. Deze worden opgesteld door datamanagement professionals of door business beleidsmedewerkers.

Aanmoedigen van veranderingsmanagement: Voor veel organisaties verschilt de 'werkwijze' binnen data governance met de huidige situatie in hen organisatie. Om data governance op te nemen in de organisatie zullen werkwijze/het gedrag van mensen moeten veranderen. Verandering is niet altijd even makkelijk en daarom dat veranderingsmanagement aangemoedigd moet worden.

Aan probleem management doen: Issue management is het proces voor het identificeren, kwantificeren, prioriteren en oplossen van data governance-gerelateerd problemen, waaronder vragen met betrekking tot rechten en procedures, problemen die komen door veranderingsmanagementprocessen, problemen omtrent het voldoen aan compliance-eisen, privacy- en vertrouwelijkheidsproblemen en problemen gerelateerd aan policies, standaarden, architectuur en procedures.

Beoordelen van Wettelijke nalegingsvereisten: Elke organisatie wordt beïnvloed door overheids- en sectorvoorschriften, inclusief voorschriften die bepalen hoe data en informatie moet worden beheerd. Een deel van de van data governance heeft te maken met het bewaken en waarborgen en naleving van de regelgeving. Naleving van regelgeving is vaak de eerste reden voor het implementeren van data governance.

Implementeren van data governance: Data governance kan niet van de ene op de andere dag worden geïmplementeerd. Het vereist een planning - niet alleen om verantwoording af te leggen over verandering, maar ook simpelweg omdat het veel complexe activiteiten omvat die gecoördineerd moeten worden. Implementatie-roadmap die tijdschema's voor, en relaties tussen verschillende activiteiten schetst is hierbij dan handig.

Ontwikkel een business glossary: Een glossary (woordenlijst) is nodig omdat mensen woorden anders gebruiken. Het is vooral belangrijk om duidelijkheid te hebben over definities van de data. Een glossary helpt hier dan bij.

Werk samen met architectuur groepen: Het business datamodel moet gezamenlijk worden ontwikkeld en onderhouden door data-architecten en Data Stewards. Afhankelijk van de organisatie kan dit werk worden gecoördineerd door de Enterprise Data Architect of door de steward.

Sponsor Data Asset Valuation; De waarde van data moet gemeten worden. Er zou een standaard moeten komen waaraan data moet voldoen.

Integreer data governance: Het is handig om een plan te hebben met betrekking tot de uitvoering van data governance. Hierin zou een lijst moeten opgenomen worden waarin alle activiteiten zijn weergegeven die nodig zijn voor het implementeren en uitvoeren van data governance activiteiten. Het schetst activiteiten, timing en technieken die nodig zijn om succesvol te zijn. Daarnaast is het ook handig om een data governance community of interest op te zetten. Vooral voor het eerste jaar is dit handig. Mensen kunnen zo beter weten wat het precies inhoudt en van elkaar leren.

Om uiteindelijk data governance te implementeren raad DAMA DMBOK aan om een onlineprogramma te gebruiken. Dit zorgt ervoor dat alle document op een centrale plek staan. Daarnaast brengen websites nog handige voordelen met zich mee. Zo kunnen bij websites documentation libraries aanwezig zijn en kunnen dingen snel opgezocht worden.

Data architecture: Een architectuur beschrijft het systeem binnen een organisatie. Het weergeeft alle componenten binnen het systeem, de relaties tot elkaar en tot de omgeving en de principes die het ontwerp en de verandering bepaald. Het is belangrijk om een goede architectuur te hebben. Dit helpt namelijk bij het begrijpen van de huidige systemen, naleving van regelgeving en maakt het effectiever. Data architectuur wordt ook wel gezien als de brug tussen de business strategie en de technologische uitvoering. Het stemt business en IT op elkaar af. Zo worden bijvoorbeeld eisen vanuit de business vertaald in data en systeem requirements om zo ervoor te zorgen dat bedrijfsprocessen uiteindelijk de juiste data binnen krijgen. Daarnaast helpt het ook bij het beheren van complexe gegevens- en informatielevering door de gehele organisatie heen.

Een data architectuur kent eigenlijk vier verschillende domeinen: Enterprise business Architecture, Enterprise Data Architect, Enterprise Applications Architecture en Enterprise Technology Architecture. Het is belangrijk dat hierbij domeinen (en dus de architecten) gezamenlijk kijken naar ontwikkelingen, richtingen en vereisten. Elk domein heeft namelijk op elkaar. In onderstaande tabel is er per domein gekeken naar het doel, het element waaruit het bestaat, afhankelijkheden en de rollen binnen zo'n domein.

<i>Domein</i>	<i>Enterprise business Architecture</i>	<i>Enterprise Data Architect</i>	<i>Enterprise Applications Architecture</i>	<i>Enterprise Technology Architecture</i>
<i>Doel</i>	Om te identificeren hoe een Organisatie waarde creëert voor klanten en andere belanghebbenden	Om te beschrijven hoe gegevens moeten worden georganiseerd en beheerd	Om de structuur en functionaliteit van applicatie in een onderneming te beschrijven	Om te omschrijven wat (fysieke) technologie nodig heeft om systemen op goede wijze te laten

	functioneren en zo waarde levert.			
<i>Onderdelen</i>	Bedrijfsmodellen, processen, mogelijkheden, diensten, evenementen, strategieën, vocabulaire	Data modellen, definitie van gegevens, data flows, data Mapping Specification gestructureerde data API's	Bedrijfssystemen, softwarepakketten, databases	Technologische platformen, netwerken, beveiliging, integratietools
<i>Afhankelijkheden</i>	Stelt requirements voor andere domeinen vast	Beheert data gemaakt en vereist door de business architecture.	Handelt op specifieke data volgens de business requirements.	Hosten en uitvoeren van Application Architecture
<i>rollen</i>	Business architecten, business analisten, business data stewards	Data architecten, data steward en modellers	Applicatie architecten	Infrastructuur architecten

Tabel 14: Enterprise architectuur domeinen (DMBOK2, 2017)

Dit kent twee activiteiten:

Vaststellen van data architectuur: Een data architectuur zou een integraal onderdeel moeten zijn van de Enterprise architectuur. Hierbij wordt er een framework gekozen dat van toepassing is op het bedrijfstype (gebruik bijvoorbeeld een overheid framework voor een overheidsinstantie). De opvattingen in het framework moeten bruikbaar zijn in de communicatie naar de verschillende belanghebbenden. Dit is vooral belangrijk voor Data Architecture-initiatieven, omdat ze betrekking hebben op bedrijven en systemen. Data-architectuur heeft een nauwe samenhang met bedrijfsarchitectuur

Integreren met Enterprise Architecture: Om Enterprise Data Architecture specificaties te ontwikkelen moet het eerst van een ontwerpniveau naar een meer gedetailleerd niveau dat in relatie is met de andere architectuur domeinen. Dit gebeurt meestal door middel van gefinancierde projecten. Gefinancierd projecten sturen over het algemeen architectuurprioriteiten aan. Toch moeten ook bedrijfswijde data architecture kwesties aangepakt worden. Data architecture kan zelf de scope van een project beïnvloeden. Het is daarom handig om Enterprise Data Architecture-kwestie te integreren met project portfolio management. Hierdoor kan een roadmap geïntegreerd worden en draagt het bij aan een beter projectresultaat. Andersom moet ook de Enterprise data architecten betrokken worden bij de ontwikkeling integratie van Enterprise application.

3.2.3 Voordelen

- Zeer theoretisch maar gaat ook in op de praktische kant
- Niet geheel onbekend voor de organisatie
- Als alle activiteiten uitgevoerd worden komen er veel documenten tot stand.
- Heel erg gericht op verandermanagement
- Voor iedereen beschikbaar
- Een zeer bekend framework voor datamanagement, bekender dan CMDC (Steenbeek, 2021)

3.2.4 Nadelen

- Het is mogelijk te uitgebreid. Er is nu naar twee kennisgebieden gekeken, maar dit framework kent er veel meer.

3.3 Data Management Capabilities

3.3.1 Waarom interessant voor Rijkswaterstaat

Rijkswaterstaat kent drie soorten databasesystemen: Oracle-database, Postgres en een MS SQL server (Azure). De reden dat ze drie verschillende databasesystemen hebben is omdat elk systeem verschillende (integratie) eigenschappen kent. Een MS SQL server past het best bij een Microsoft omgeving. Andere systemen kunnen hier ook mee werken, maar dit gaat niet heel makkelijk. Om deze reden hebben ze dus 3 verschillende databasesystemen (Rijkswaterstaat, 2018). Zo'n systeem kent verschillende databases en binnen/tussen de databases kan data verplaats, gecontroleerd en beheerd worden. Het Clouw Data Mangement Capabilities (CMDC) Framework gaat hier ook voornamelijk op in en kijkt hoe dit goed gedaan kan worden tussen verschillende cloud omgevingen.

3.3.2 Wat houdt het in

Het Clouw Data Mangement Capabilities (CMDC) Framework definieert de best practice mogelijkheden die nodig zijn om data te managen en te controleren in een Cloud omgeving. Het doel van dit framework is om onder andere vertrouwen en betrouwbaarheid op te bouwen in Cloud technologie. CMDC is een best practice assessment en framework voor het manage en controleren van data in single, multiple en hybride Cloud omgevingen. Het wordt gebruikt om de capaciteiten van een organisatie, die nodig zijn voor het ondersteunen van een gecontroleerde integratie en migratie naar de Cloud omgeving te beoordelen. Het focust zich op, en bereidt zich uit op de mogelijkheden die cruciaal zijn voor het controleren van belangrijke en gevoelige data, en benadrukt kenmerken van hedendaags Cloud platforms die kansen bieden voor standaardisatie en automatisering van databeheer en -controle. (EDM Council Inc, 2021)

Om data in een Cloud omgeving te beheren zijn er een paar datamanagement capabilities:

- Een organisatie moet duidelijke de verantwoording, beheer en governance vaststellen voor data die zijn gemigreerd naar of gemaakt in Cloudomgevingen
- Het is een vereiste dat men te allen tijde weet wat voor data er in de Cloud omgeving zit en hoe gevoelig deze data is. Dit bijhouden is essentieel om automatisch beheer van datagebruik en data toegang.
- Datamanagement beheer moet vastgesteld worden aan de hand van de levenscyclus van de data
- Data assets moeten geschikt zijn voor de beoogde doelen en zich houden aan de eisen van bewaring en archivering.
- Net als bij on-premises data-assets, is het ontwerp van de data-architectuur en de configuratie van ondersteunende technologieën belangrijk om ervoor te zorgen dat de bedrijfsdoelstellingen worden gehaald. (EDM Council Inc, 2021)

CDMC legt de vereisten voor deze capabilities vast in zes gebieden. Deze zes componenten van het frameworks, bevatten in totaal 14 capabilities en 37 sub- capabilities. Dit framework kent de volgende zes gebieden: Governance & Accountability, Cataloging & Classification, Accessibiliy & Usage, Pretectioin & Privacy, Data lifecycle en Data & Technical Architecture. Om binnen de scope te blijven wordt nu alleen gekeken naar Governance & Accountability en Data & Technical Architecture.

Governance & Accountability: Dit gebied vormt de basis voor succesvol datamanagement in Cloudomgevingen. De Cloudomgeving introduceert uitdagingen en kansen voor schaalvergroting, standaardisatie, automatisering en het gedeelde verantwoordelijkheidsmodel. Daarom is het belangrijk om een effectief data governance programma toe te passen op data

die zich in een Cloud omgeving bevindt. Alle belanghebbenden moeten een duidelijk begrip hebben van gegevenscontroles en verantwoordelijkheid voor elke rol. De aanpak is vergelijkbaar met hoe data governance, -controls en -accountability zijn toegepast op gebruikelijk datamanagement in een organisatie.

Hierbij is de scope als volgt:

- Definieer business cases voor het managen van data in de Cloud.
- Zorg ervoor dat rollen en verantwoordelijkheden van data-eigenaren zich ook uitbreiden naar data in de Cloud.
- Zorg ervoor dat databronnen beheerd wordt aan de hand van goede bronnen en distributeurs.
- Automatisering van de Cloud zorgt voor kansen met betrekking tot data consumption.
- Begrijpen van requirements voor het beheer van datasoevereiniteit en de risico's van cross-border data movement.
- Implementeer controles voor datasoevereiniteit en cross-border data movement (EDM Council Inc, 2021)

(Datasoevereiniteit heeft te maken met de data van de burgers in het land. De overheid probeert te voorkomen dat deze gegevens in de verkeerde handen vallen, dit doet zij door maatregelen die ervoor moeten zorgen dat bedrijven persoonsgegevens niet in het buitenland vrij laten.)

Data & Technical Architecture: Data & Technical Architecture gaat over de architectonische problemen die uniek zijn voor de Cloud en hoe dit het manage van data in de Cloud omgeving beïnvloed. De meeste Cloud serviceproviders bieden verschillende opties aan voor hoe business oplossingen middels verschillende Clouds service ontworpen en geïmplementeerd kunnen worden. Het ontwikkelen en toepassen van bepaalde architectuurpatronen en richtlijnen kunnen zorgen voor een basis met betrekking tot data management in de Cloud. (EDM Council Inc, 2021)

Dit gebied bevat een aantal capabilities om ervoor te zorgen datastromen in, uit of tussen Cloudomgevingen begrepen worden en dat architectuur richtlijnen helpen bij het ontwerpen ontwerp van Cloud computing-oplossingen.

Het gebied focust op de volgende dingen:

- Het vaststellen en toepassen van principes voor beschikbaarheid van data.
- Ondersteunen van business requirements voor backups en herstellen van data
- Optimaliseren van het gebruik van de Cloud en de bijbehorende kosten voor de service.
- Ondersteunen van data portability en de mogelijkheid om data tussen verschillende Cloud omgevingen te verplaatsen
- Automatiseren van identificeren van dataprocessen en datastromen in de Cloud omgeving zelf en tussen verschillende Cloud omgevingen.
- Het identificeren, volgen en managen van verandering op gebied van oorsprong van de data. Dit maakt het mogelijk om de oorsprong van data altijd te kunnen achterhalen en uit te leggen.
- Tools die op zinnige wijze rapporteren en visualiseren, zowel vanuit een business perspectief als een technisch perspectief (EDM Council Inc, 2021)

3.3.3

Voordelen

- Heel erg cloud gericht, meer dan DAMA DMBOK
- Gaat vooral in op hoe data binnen en tussen cloudomgevingen gemanaged en gecontroleerd kan worden.
- Praktische focus

3.3.4

Nadelen

- Wordt verwacht dat de organisatie een sterke basis heeft op gebied van datamanagement
- Zeer onbekend voor de organisatie
- De kijk op implementatie van datamanagement ontbreekt
- Afgeleid van DCAM, DCAM is niet toegankelijk voor iedereen. Daarom dat ze dit hebben gepubliceerd. Hierin wordt voornamelijk de aanpak besproken voor dit framework.
- Om volledige toegang te krijgen tot dit framework is een organisatie lid worden van the EDM council, dit is redelijk prijzig.

3.4

Conclusie

Er is dus gekeken naar twee verschillende frameworks; DAMA DMBOK en Cloud Data Management Capabilities Framework. Het advies hierbij is om vooral het DAMA DMBOK framework aan te houden. Hier zijn meerdere redenen voor. Ten eerste wordt bij het Cloud Data Management Capabilities Framework al verwacht dat de organisatie een sterke basis heeft op gebied van datamanagement. Dit

framework gaat daar namelijk nog veel dieper op in. De vraag is dan, is op het moment datamanagement binnen Rijkswaterstaat al zover, heeft het al zo'n sterke basis? Binnen Rijkswaterstaat zijn er een paar documenten aanwezig over data governance, maar de datastrategie is ook nog maar twee jaar oud. Mogelijk dat Rijkswaterstaat nog niet op het niveau zit om dit framework toe te passen. Daarnaast is er niet veel theorie omschreven in dit framework. Het heeft meer een praktische focus. Om die reden kan het ook moeilijker zijn voor beginners om dit te begrijpen.

DAMA DMBOK is echter weer heel erg theoretisch. Het verschaft veel informatie over datamanagementcomponenten, maar heeft ook een praktische kant. Heeft het ook over ervaringen en laat zien wat echt werkt. Een andere reden waarom DAMA DMBOK aan te raden is, is omdat dit in de organisatie al een beetje bekend is. In ieder geval meer bekend dan Cloud Data Management Capabilities Framework. Voordat ik kwam was er al aangegeven dat dit een goed framework was om te gebruiken, dit impliceert dat op z'n minst een paar mensen binnen de organisatie kennis hebben van DAMA DMBOK. Het is moeilijk om zaken binnen een grote organisatie als Rijkswaterstaat te veranderen. Er was namelijk al via de interviews naar voren gekomen dat zaken pas veranderen, als dit voor een probleem begint te zorgen. Het is daarom misschien ook niet verstandig om een geheel nieuw framework te introduceren, terwijl ze al een ander prima bruikbaar framework kennen.

Als er gekeken wordt naar de requirements en deze in een L-matrix uitgedrukt worden dan is het volgende te zien:

Tabel 15: L-matrix requirements DAMA DMBOK vs CDMC

<i>Requirement</i>	<i>DAMA DMBOK</i>	<i>Score</i>	<i>CDMC framework</i>	<i>Score</i>
<i>De governance rollen worden ingevuld</i>	Duidelijk, een van de eerste activiteiten binnen DAMA DMBOK – governance is het definiëren van wat er wordt beheerd en wie dit beheerd.	5	Matig, Het gaat meer over hoe rollen en verantwoordelijkheden van data-eigenaren zich ook uitbereid naar data in de Cloud	3
<i>Er is een erkende eigenaar voor de data aangewezen</i>	Duidelijk, dit framework gaat ook in op data ownership. Zo wordt er onderscheid gemaakt tussen verschillende data owners.	5	Duidelijk, een van de doelen is ook om rollen van data eigenaren te definiëren en dit uiteindelijk ook toe te passen in de cloud omgeving.	5
Totaal		10		8

Hierbij geldt: 5: voldoet in hoge mate, 3 Voldoet in zekere mate, 1, is enigszins van toepassing. (managementmodellensite, 2021)

Kortom, kijkend naar de organisatie, de requirements van de organisatie en hoever de organisatie is met data management is het aan te raden om vooral te kijken naar het DAMA DMBOK.