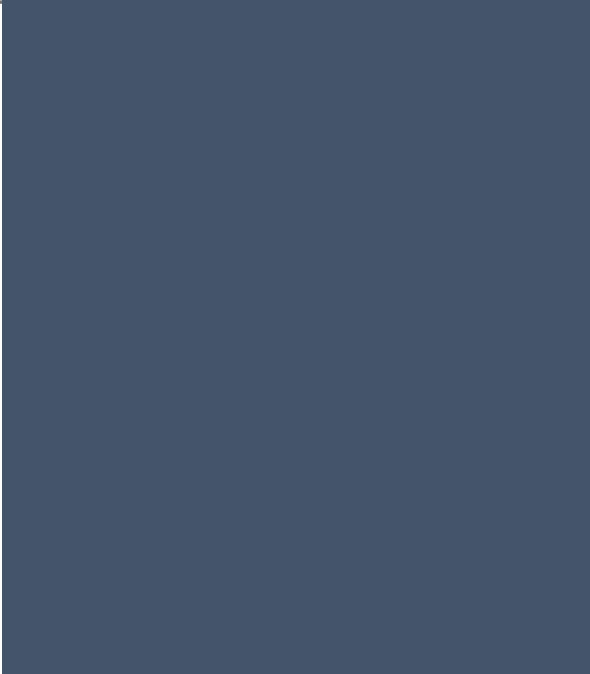




De bruikbaarheid van metadata bij de NS

Stageverslag

Naam: Glenn de la Rie
Stagebedrijf: NS-DIA
Studie: Informatica Hogeschool Leiden
Specialisatie: BDAM
Versie Type: final versie
Versie nummer: 4.0
Versie Datum: 13-6-2022



Voorwoord

Voor u ligt de scriptie 'Bruikbaarheid van metadata bij de NS'. Deze opdracht is uitgevoerd bij het Enterprise datamanagement team van de NS. Deze scriptie is geschreven ter afronding van mijn opleiding Informatica aan de Hogeschool Leiden. Ik ben bezig geweest met deze opdracht van februari 2022 tot en met juni 2022.

Ik ben per toeval bij het EDM-team terecht gekomen, nadat ik had verloren van Tim met een Mario kart toernooi op het kantoor van de NS. Zo raakte wij aan de praat en daaruit is mijn stage voortgekomen.

Mijn samenwerking met mijn stagebegeleider Tim Reacq en opdrachtgever Gert Jonkman is prettig verlopen. Samen met hen heb ik mijn onderzoeksvragen opgesteld en ik kon altijd bij ze terecht voor vragen en het doorlezen van mijn documenten. Hiervoor wil ik ze graag bedanken. Ook wil ik ze bedanken voor de vrijheid om zelfstandig aan mijn project te werken en mijn ideeën in te brengen.

Tijdens mijn afstuderen heb ik veel geleerd en kennis gemaakt met de NS als werkgever. Ik wil de medewerkers van het EDM-team bedanken voor de mogelijkheid om de wekelijkse bijeenkomsten bij te wonen. En voor de kans om te mogen solliciteren voor de functie als Cloud Datamanagement ontwikkelaar.

Tot slot wil ik mijn vriendin en ouders bedanken voor hun steun en aanmoediging tijdens dit proces.

Glenn de la Rie

Leiden 14-06-2022

Samenvatting

De NS ziet dat er minder gebruik wordt gemaakt van metadata dan ze verwachten. Mogelijk komt dit door de gebruiksvriendelijkheid van het huidige metadata portaal. Het doel van deze scriptie is dan ook om te onderzoeken wat de NS kan doen om het beste te kunnen halen uit hun metadata.

Ik heb kwalitatief onderzoek gedaan naar de best-practices van metadata management door literatuur te raadplegen. Daarnaast heb ik het probleem verhelderd m.b.t. het huidige metadata portaal van de NS en heb ik een testmethode ontwikkeld voor het meten van de bruikbaarheid hiervan.

Uit mijn onderzoek komt naar voren dat er onvoldoende informatie beschikbaar is om het portaal te laten bepalen welke metadata relevant is voor de gebruiker. Wel zijn er mogelijkheden om de gebruiker zelf zijn context te laten bepalen.

In het huidige metadata portaal is nog ruimte om de metadata te verduidelijken. Het metadata portaal moet de gebruiker kunnen leren wát metadata is zodat ze het portaal kunnen gebruiken.

Daarnaast is het doel van een metadata portaal om gebruikers zo snel mogelijk hun metadata te laten vinden. Door het toevoegen van extra functionaliteiten die bijdragen aan de diverse use cases kan iedere gebruiker gemakkelijk vinden wat hij nodig heeft. Een belangrijke factor om rekening mee te houden is dat niet elke gebruiker evenveel technische kennis heeft.

Mijn aanbeveling is om verder onderzoek te doen naar wat verschillende gebruikers/groepen binnen de NS precies willen halen uit de metadata.

Inhoud

Voorwoord	1
Samenvatting.....	2
1 Begrippenlijst.....	5
2 Inleiding	6
3 Organisatie	7
3.1 NS.....	7
3.2 Stageafdeling	7
3.3 Hoe is de afdeling ingericht?	8
3.4 Stagebegeleiding	8
4 Opdracht.....	9
4.1 Aanleiding.....	9
4.2 Doelstelling.....	9
4.2.1 Bedrijfsdoelstelling.....	9
4.3 Oplevering	10
4.4 Projectgrenzen	11
4.4.1 Scope	11
4.5 Onderzoek	12
4.5.1 Hoofdvraag	12
4.5.2 Deelvragen.....	12
4.6 Technische aspecten	13
5 Planning.....	14
6 Methodiek	15
7 Onderzoek	16
7.1 Hoe wordt datamanagement binnen de NS toegepast?.....	16
7.2 Hoe wordt datamanagement informatie beschikbaar gesteld voor de organisatie?	19
7.3 Wat zijn de belangrijkste gebruikersgroepen van datamanagement informatie?	21
7.4 Waardoor is het gebruik van EDM metadata minder dan de potentiële mogelijkheden? ...	23
7.5 Wat zijn wensen van medewerkers over het gebruik en toegankelijkheid van metadata? .	24
7.6 Wat zijn de huidige best practices van metadatamanagement?	24
7.7 Hoe kan de bruikbaarheid (relevant, eenduidig en gebruiksvriendelijk) van het EDM metadata portaal worden gemeten?	25
8 Proof-of-concept	26
8.1 Functionaliteiten	26
8.2 Schets structuur.....	27

8.3	Data	28
8.4	Designs	29
8.5	API.....	29
8.6	Front-end.....	31
9	Conclusie	32
10	Aanbevelingen.....	33
11	Discussie	34
12	Reflectie.....	35
12.1	Stageproces	35
12.2	Product	36
12.2.1	Requirements analyse	36
12.3	Persoonlijke reflectie.....	36
13	Competenties	37
13.1	A-competenties	37
13.1.1	Requirements analyse	38
13.2	B-competenties	39
14	Bibliografie.....	40

1 Begrippenlijst

Begrip	Definitie
Enterprise datamanagement	Is het geheel van activiteiten binnen organisaties dat zich richt op het op een gestructureerde wijze identificeren, classificeren, registreren, modelleren, ontsluiten, beveiligen, archiveren en vernietigen van gegevens.
ETL-proces	De processen die gebruikt worden om gegevens uit verschillende databases te verenigen in een andere database, veelal een datawarehouse.
Metadata	Metadata zijn alle gegevens over de data zelf, zoals eigenaarschap, wijzigingshistorie, BIV-classificatie etc.
Metadatamanagement	Beschrijft het proactieve gebruik van metagegevens om gegevens te beheren op een manier die uiteindelijk goed geïnformeerde zakelijke beslissingen en efficiëntie ondersteunt.
Product owner	Is verantwoordelijk voor het maximaliseren van de waarde van zowel het product als het projectteam. Hoe dit precies gedaan wordt, verschilt vaak per organisatie
Proof of concept	Is een realisatie van een bepaalde methode of idee om de haalbaarheid ervan aan te tonen, of een demonstratie in principe met als doel te verifiëren dat een concept of theorie praktisch potentieel heeft
System of record	Is een term voor gegevensbeheer voor een informatieopslagsysteem dat de gezaghebbende gegevensbron is voor een bepaald gegevenselement of stuk informatie.
System of engagement	Wordt gebruikt om te communiceren met de gebruikers, mogelijk door gebruik te maken van een system of record.

2 Inleiding

Veel organisaties hebben ‘meer data gedreven werken’ als doel binnen hun bedrijf. De NS wil ook graag meer gebruik gaan maken van de data die ze hebben omdat dit bijdraagt aan het maken van beslissingen en behoud van kennis. Naast het hebben van een goede infrastructuur en een passende bedrijfscultuur is Enterprise Data Management een belangrijke factor om de transitie te maken naar een data gedreven organisatie. Metadata zijn essentieel om gebruik te kunnen maken van data.

Dit verslag in opdracht van de NS gaat over het organisatie breed beschikbaar maken van metadata. Het eerste deel van het verslag is om context te bieden aan de opdracht; toelichting van de organisatie, evenals het doel van het project en deliverables en het te verrichten onderzoek.

In het tweede deel zijn de resultaten en de vervolgstappen naar het eindproduct verwerkt. Ook het advies voor de opdrachtgever is hier terug te vinden.

Het verslag wordt afgesloten met een reflectie op de afstudeerstage en persoonlijke ontwikkeling.

3 Organisatie

De afstudeeropdracht is uitgevoerd in opdracht van de NS. In dit hoofdstuk wordt de organisatie beschreven en er wordt ook ingegaan op de plaats van de stageafdeling in de organisatie. Het doel van dit hoofdstuk is om context te bieden over mijn stageplek.

3.1 NS

De NS is een grote onderneming met een belangrijke maatschappelijke functie in de Nederlandse samenleving met de Nederlandse staat als enige aandeelhouder. De NS hecht veel belang aan openheid en transparantie. De NS is actief in de wereld van openbaar vervoer. Ze bevorderen het gebruik van het openbaar vervoer en zorgen ervoor dat iedereen op een prettige manier kan reizen. De NS heeft meer dan 400 stations. Het hoofdkantoor staat in Utrecht. De NS heeft in totaal meer dan 20.000 medewerkers waarvan meer dan 1000 eigen IT'ers.

Voorheen was de NS een organisatie die was opgesplitst in zelfstandige units. Ieder bedrijfsonderdeel had zijn eigen IT-afdeling en werkte los van elkaar. Tijdens het implementeren van de OV-chipkaart 12 jaar geleden heeft de NS besloten dat het bedrijf één geheel moet worden waarin Business en IT samen over de gehele breedte en diepte intensief en data gedreven samenwerken. Dit is een grote verandering geweest en heeft geleid tot herinrichting van NS als geheel en de afdeling DIA (Data innovatie analyse) in het bijzonder. De afdeling DIA bestaat uit ruim 450 medewerkers. Enterprise Datamanagement (EDM) is hier een onderdeel van. Er werken hier ongeveer 45 werknemers.

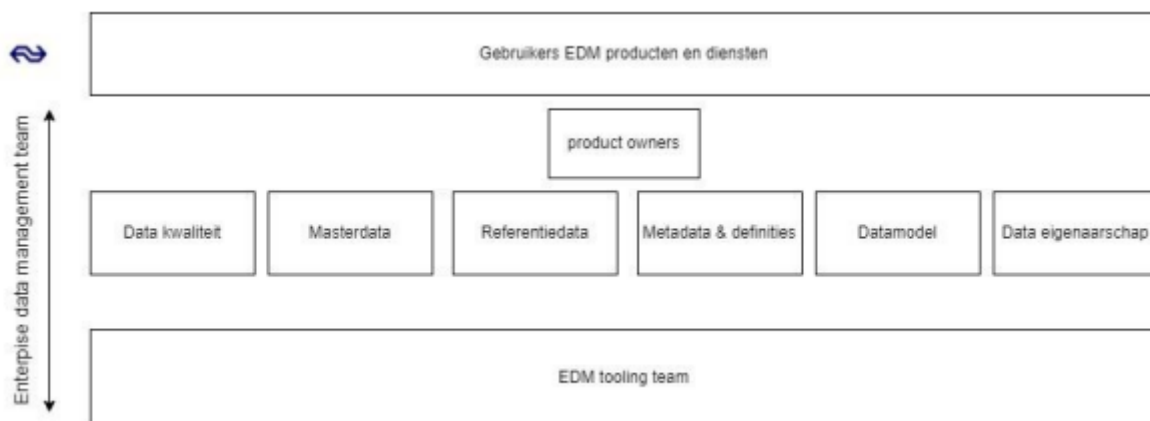
3.2 Stageafdeling

Ik ga mijn afstudeerstage lopen bij de afdeling Enterprise Data Management (EDM).

Datamanagement is het zorgen voor standaardisatie en herbruikbaarheid van data over de hele NS-organisatie. De datamanagement afdeling is gesplitst in een aantal onderdelen geïnspireerd uit het DAMA DMBOK wiel (zie Bijlage 1 EDM in context). Elk onderdeel binnen de afdeling heeft zijn eigen systemen voor het opslaan van alle (meta)data. Deze afdeling is nu bezig alle (meta)data in één systeem te krijgen. Daarvoor gebruiken ze de tools van het bedrijf Informatica.

3.3 Hoe is de afdeling ingericht?

Hieronder is een vereenvoudigd beeld van de EDM afdeling weergegeven. Alle datamanagement onderdelen worden ondersteund door het tooling team. Ook heeft elk onderdeel inclusief het tooling team zijn eigen product owners. En er zijn nog extra product owners die vooral gericht zijn op de gebruikers van EDM-producten en diensten. Deze externe product owners achterhalen de wensen en eisen van gebruikers en communiceren dit met de EDM-teams.



Figuur 1 EDM-team

3.4 Stagebegeleiding

De begeleider tijdens mijn afstudeerstage is Tim Reacq. De opdrachtgever is Gert Jonkman. Beide begeleiden mij tijdens dit project. In de rest van het verslag noem ik ze 'stagebegeleiders'.

Daarnaast heb ik contact met het EDM-team. Tim Reacq is een masterdata specialist en Gert Jonkman is de product owner van het EDM tooling team.

4 Opdracht

Dit hoofdstuk heeft als doel om inzicht te geven in de afstudeeropdracht.

4.1 Aanleiding

Het EDM-team heeft geconcludeerd dat er nog niet zoveel gebruik wordt gemaakt van metadata als werd gehoopt. De huidige manier van het weergeven van metadata is niet gebruiksvriendelijk genoeg. Met als gevolg dat medewerkers het overzicht kwijt zijn en de informatie niet kunnen vinden en gebruiken. Terwijl dit de medewerkers juist zou moeten helpen.

Het EDM-team heeft wel alle benodigde en juiste metadata in interne systemen met de juiste strategieën voor het omgaan met data (system-of-record) maar heeft problemen met het beschikbaar en bruikbaar maken van deze metadata naar de eindgebruikers (system-of-engagement).

4.2 Doelstelling

Het doel van dit project is om de NS te adviseren hoe metadata toegankelijker kan worden gemaakt voor medewerkers in de organisatie.

4.2.1 Bedrijfsdoelstelling

De afdeling EDM heeft als doelstelling om NS-breed de Enterprise Data te standaardiseren, te optimaliseren en te regisseren zodanig dat het voor NS en hun partners waarde toevoegt en de onderlinge communicatie vereenvoudigt. Hiervoor legt EDM de metadata, governance en definities centraal vast. Deze informatie is momenteel onvoldoende eenvoudig toegankelijk voor de gebruikers die in hun context slecht een deel van alle beschikbare informatie gebruiken.

4.3 Oplevering

Het belangrijkste eindproduct is een advies voor de NS. Het advies moet NS verschillende opties bieden die kunnen helpen om metadata beter weer te geven voor gebruikers. Daarnaast lever ik ook nog andere producten in ter ondersteuning. Deze zijn iets veranderd in vergelijking met de originele planning. Hieronder een lijst van de overige deliverables:

- Designs van de gegeven adviezen: Van alle adviezen worden wireframes gemaakt die net als het proof of concept bijdragen aan het inzichtelijk maken van de adviezen.
- Document best practices metadatamanagement.
- Softwarecomponenten: Proof-of-concepts projecten die laten zien hoe adviezen er in de praktijk uit kunnen zien.
- Data modellen: Ik maak een aantal datamodellen voor de huidige situatie en een model voor het implementeren van de proof-of-concepts.
- Test methode: Ik lever een methode op om de bruikbaarheid van een metadata portaal te testen.

4.4 Projectgrenzen

Aan het begin van het afstudeertraject is een scope gemaakt en zijn de afhankelijkheden gemaakt.

4.4.1 Scope

Voordat ik mijn stage begon heb ik een plan van aanpak gemaakt. Sommige dingen zijn echter anders gegaan dan verwacht. Hieronder beschrijf ik kort de scope met de dingen die anders zijn gegaan.

De probleemstelling van de stage is gericht op het toegankelijk maken van de data. Er hoeft dus geen nieuwe metadata te worden verzameld omdat deze al beschikbaar is.

De afdeling wil graag dat ik een advies geef zonder dat ik mij vooraf laat beïnvloeden door de beperkingen van de tooling die NS al heeft. Ik word gestimuleerd om out-of-the-box te denken en dit mee te nemen in mijn advies.

Er wordt geen nieuwe connectie gebouwd tussen alle data managementsystemen, omdat er al een traject loopt om deze systemen te vervangen door de tools van Informatica. Ik mag er van uit gaan dat dit al gereed is. Dit brengt veel nieuwe metadata soorten met zich mee die ik kan gebruiken tijdens het afstudeerproject.

In het Plan van Aanpak heb ik aangegeven dat alle metadata wordt aangeleverd door het EDM-team inclusief de nieuwe metadata soorten. Tijdens de stage bleek echter niet mogelijk voor het EDM-team om alle metadata aan te leveren. Daarom heb ik zelf test data gemaakt met een structuur die lijkt op die van de metadata in de systemen.

In de bijlage 'EDM in context' wordt de NS EDM cirkel toegelicht. De volgende onderdelen horen bij deze cirkel, maar vallen buiten mijn scope:

- Referentiedata en Masterdata: Mijn afstudeerstage is gericht op metadata, deze onderdelen vallen hier grotendeels buiten.
- Gebruik en Toegang: Deze onderdelen zijn gescheiden van de andere onderdelen. En er is geen behoefte om deze beschikbare metadata te integreren met de rest.

4.5 Onderzoek

Tijdens de eerste fase van mijn afstudeerproject ben ik voornamelijk bezig geweest met een vooronderzoek. Hierin beantwoord ik een aantal van de deelvragen en doe ik extra kennis op over de huidige data-architectuur en datamanagement.

4.5.1 Hoofdvraag

Hieronder de hoofdvraag die is opgesteld samen met de opdrachtgever. De aard van het probleem is redelijk subjectief maar de hoofdvraag is zo SMART mogelijk geformuleerd (exclusief tijdsgebonden maar dat is automatisch mijn stageperiode):

“Hoe kan de NS EDM-afdeling de aanwezige EDM-informatie beschikbaar stellen voor inzage en gebruik binnen de NS,

zodanig dat deze relevant, eenduidig en gebruiksvriendelijk binnen de context van het gebruik ter beschikking is.”

Ter verduidelijking van de vraag wil ik naar twee specifieke gedeeltes kijken.

“zodanig dat deze relevant, eenduidig en gebruiksvriendelijk”

De begrippen relevant, eenduidig en gebruiksvriendelijk kunnen als maatstaven dienen voor de bruikbaarheid van de EDM-informatie voor de gebruikers. Deze begrippen zijn gedefinieerd op basis van de vragen die in de nulmeting komen te staan.

“binnen de context van het gebruik ter beschikking is.”

Gebruikers van EDM-informatie krijgen op dit moment alle beschikbare informatie binnen. De gebruiker moet vervolgens zelf bepalen wat relevant is. Door de hoeveelheid data werkt deze methode niet. Het EDM-team wil daarom overgaan naar een oplossing waar gebruikers eerst informatie krijgen die relevant is voor de gebruiker. Vervolgens kan de gebruiker altijd nog extra informatie zoeken als dat nodig is.

4.5.2 Deelvragen

Om de hoofdvraag te beantwoorden zijn er een aantal deelvragen opgesteld. In het plan van aanpak wordt de aanpak per deelvraag beschreven:

1. Hoe wordt datamanagement binnen de NS toegepast?
2. Hoe wordt datamanagement informatie beschikbaar gesteld voor de organisatie?
3. Wat zijn de belangrijkste gebruikersgroepen van datamanagement informatie?
4. Waardoor is het gebruik van EDM metadata minder dan de potentiële mogelijkheden?
5. Wat zijn wensen van medewerkers over het gebruik en toegankelijkheid van metadata?
6. Wat zijn de huidige best practices van metadatamanagement?
7. Hoe kan de bruikbaarheid (relevant, eenduidig en gebruiksvriendelijk) van het EDM metadata portaal worden gemeten?

4.6 Technische aspecten

Ik heb ervoor gekozen om mijn advies rapport te verdedigen met verschillende proof of concepts die laten zien hoe mijn adviezen er in het echt uit zouden kunnen zien. Dit maakt mijn advies visueel en kan helpen met het kiezen van bepaalde design keuzes voor de NS in de toekomst voor een metadata portaal.

5 Planning

Mijn afstudeertraject is op te splitsen in een vier fases.

Fase 1: Bepalen stageopdracht

Tijdens de start van de stage was het belangrijk om een beter beeld te krijgen van de opdracht die ik zou gaan uitvoeren. Belangrijk was om samen met de opdrachtgever te bepalen wat wel en niet binnen de grenzen van het project hoorde. Dit is vermeld in de vorige hoofdstukken en het plan van aanpak.

Fase 2: Uitvoeren onderzoek

Nadat het PVA is goedgekeurd ging ik aan de slag met de deelvragen. In hoofdstuk 7 van dit verslag geef ik antwoord op deze vragen. Per deelvraag is ook een externe bijlage waar deze uitgebreider wordt behandeld.

Tijdens deze fase heb ik meer geleerd over data management en vooral metadatamanagement. Omdat ik later in de stage de proof of concepts zou gaan maken heb ik ook twee cursussen gedaan over Vue (front-end framework) en typescript.

Aan het eind van het onderzoek maak ik een advies zie externe bijlage 'advies'. De volgende fases zijn gebaseerd op het advies dat ik heb gegeven voor het verbeteren van het metadata portaal.

Fase 3: ontwerpen en implementeren backend

Nu ik genoeg achtergrond informatie heb en de adviezen kan ik beginnen met het ontwerpen en implementeren van de backend waar de proof-of-concepts gebruik van gaan maken.

Dit is begonnen met het modeleren van de huidige data en het modelleren van de data in de wenselijke situatie. Dit heb ik gedaan door conceptuele en technische data modellen te maken. Voor meer informatie zie Bijlage 'TDM'.

Zodra deze klaar was is de database gemaakt en een REST-API omheen gebouwd. De font-end kan nu makkelijk communiceren met de database.

Fase 4: Maken van de proof-of-concepts

In de laatste fase van mijn stagetraject ben ik vooral bezig geweest met het maken van de designs voor een nieuw metadata portaal en met het implementeren van een aantal van deze designs in de vorm van een proof-of-concept.

6 Methodiek

Tijdens het uitvoeren van mijn stage heb ik de KanBan methode toegepast (Lean Six Sigma Groep, sd). Door gebruik te maken van KanBan kan ik makkelijk op veranderingen inspelen. KanBan is ook een goede methode omdat ik dit project alleen doe. Het is niet nodig om dagelijkse stand-ups te doen.

Tijdens mijn stage heb ik regelmatig contact gehad met mijn begeleiders om te zorgen dat het eindresultaat in lijn is met wat de opdrachtgever wil. Elke week was er via teams een feedback moment. Daarnaast kon ik altijd vragen stellen via de mail

Tijdens mijn afstudeertraject heb ik gebruik gemaakt van verschillende onderzoeksmethodes. Het onderzoek dat ik heb gedaan was een kwalitatief onderzoek. De methodes die zijn toegepast worden hieronder weergegeven.

Onderzoeksmethode	Wat	Toegepast bij
Literatuuronderzoek	Wat zegt de literatuur over bepaalde onderwerpen	Deelvraag 1,3, 6, 7
Peer review	Feedback van stakeholders	Deelvraag 1,2,4, 5, 7
Document analyse	EDM heeft veel documenten die kunnen worden bekeken om context te krijgen	Deelvraag 1,3
Requirements analyse	Brengt de eisen en wensen in kaart	Deelvraag 5,
Stakeholder analyse	Hierbij breng je de belanghebbenden en hun belangen in beeld	Deelvraag 3

7 Onderzoek

In dit hoofdstuk ga ik verder in op het onderzoek dat ik heb uitgevoerd en behandel ik de deelvragen.

7.1 Hoe wordt datamanagement binnen de NS toegepast?

Aanpak

Het doel van deze deelvraag is het geven van context. Om deze vraag te beantwoorden heb ik naar beschikbare documenten van het EDM-team gekeken en heb ik veel gesprekken gehad met mijn begeleiders die beide een uitgebreide kennis hebben van datamanagement.

Alle verzamelde informatie staat in de externe bijlage 'EDM in context'.

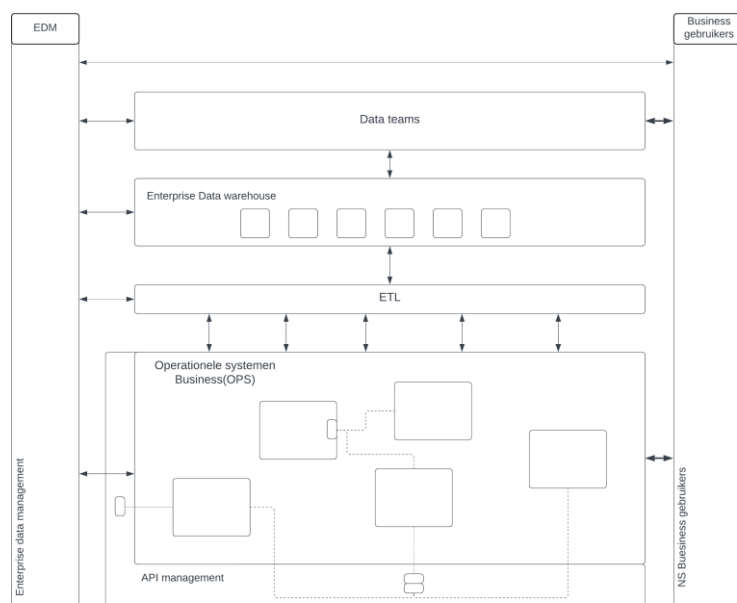
Resultaat

Het doel van EDM is de NS in staat stellen versneld te digitaliseren door data hiervoor geschikt te maken en aantoonbaar onder controle te zijn met focus op het business belang. De waarden die EDM toevoegt aan de NS zijn opgesplitst in een aantal onderdelen:

- Eenduidigheid
- Snelheid
- Betere besluitvorming
- Betrouwbaar NS

EDM heeft invloed op alle onderdelen in de organisatie omdat het elke stap van het dataproces wil standaardiseren. EDM zit midden in een traject om de huidige tooling over te zetten naar die van Informatica zodat het extra tools kan gebruiken die inzicht kunnen geven in het dataproces.

Hiernaast staat een versimpelde weergave van EDM bij de NS. Dit helpt om een overzicht te krijgen met welke gebieden EDM zich bezighoudt.



Figuur 2 EDM bij de NS

Het EDM-team heeft zijn eigen indeling van datamanagementgebieden, te zien in figuur 2. In afbeelding 3 is te zien welke taken horen bij elk datamanagementgebied.



Figuur 3 EDM cirkel

Vanuit de EDM Cirkel naar EDM Services

EDM Cirkel	(K)PI definities	IP- Eigenaarschap en IP- metadata	Datakwaliteit	Masterdata	Referentiedata	Metadata & Definities	Data-Model en catalog	Data Eigenaarschap	Toegang	Datagebruik
EDM PO (kernteamlid)	Hans	Madeleine	Madeleine	Ben (Hans)	Myrna	Chantal (Myrna)	Madeleine	Myrna	Susan	Hans
EDM Capabilities	Definitie management	IP governance management	Datakwaliteit management	Masterdata management	Referentiedata management	Metadata management	Data Modelling management	Data Governance management	User en Autorisatie management	(Secundair) Datagebruik management
EDM Services product en werkwoord	Definitiebeheer sturen	Rollen mbt IP onderhouden	DQ data profilen	Stamdatabeheer uitvoeren	Beheer van referentiedata sturen	Metadata beheer voor data-entiteiten sturen	Het opstellen van logisch datamodellen sturen	Rollen mbt datagovernance onderhouden	User / Autorisatie beheer finance systemen uitvoeren	Secundair datagebruik beoordelen (DUB proces)
	Definitiebeheer faciliteren	Rollen mbt IP beleggen	DQ eisen vastleggen		Beheer van referentiedata faciliteren (TAB5)	Metadata beheer voor data-entiteiten faciliteren	Opstellen logisch datamodellen faciliteren	Rollen data-eigenaren en stewards beleggen	User beheer DIA uitvoeren	DUR register beheren en ontsluiten
	Definities ontsluiten	Vastlegging van metadata bij IP's sturen	DQ eisen en scores raadplegen		Referentiedata ontsluiten	Metadata over data-entiteiten en business objecten ontsluiten	Synchronisatie LDM met BOM en Metadata coördineren	Datadomein-landkaart onderhouden en ontsluiten	Role Baseded row level Access inrichten voor DIA Informatie producten	Beleed externe transparantie opstellen en onderhouden
		Vastlegging van metadata bij IP's faciliteren					Data Catalog faciliteren	Data-eigenaren en stewards opleiden en begeleiden	RBAC NS breed inrichten en beheren	Voor datagebruik relevant beleid en kaders verzamelen en ontsluiten
		Metadata bij IP's ontsluiten					Data Catalog ontsluiten		RBAC NS breed uitvoeren	
		Data lineage faciliteren								
		Data lineage ontsluiten								

xxx sturen = Kaders maken voor het opstellen en beheren van xxx en naleving hiervan afdwingen
 xxx faciliteren = infrastructuur aanbieden voor beheer van xxx
 xxx ontsluiten = vastgelegde gegevens toegankelijk maken voor de gewenste doelgroep(en)

Figuur 4 taken van EDM

Hieronder staan de informatiebronnen die het EDM beheert of tot zijn beschikking heeft. Ik heb de informatiebronnen verdeeld in huidige aanwezige bronnen en toekomstige bronnen:

- Huidig
 - Data-domeinen en eigenaren lijst
 - Definities
 - BOM
 - Business object + metadata
 - Datadomein kaart
 - Logische datamodellen
 - Data scorecards en regels
 - Referentie en Masterdata
 - DUR register
- Toekomst
 - Technische metadata EDW en ETL
 - Informatie Product eigenaren en metadata
 - K(PI) definities
 - Power BI metadata

Voor meer uitleg zie de bijlage 'EDM in context'. Deze bijlage bevat ook informatie over de data en systemen die worden beheerd door het EDM-team.

7.2 Hoe wordt datamanagement informatie beschikbaar gesteld voor de organisatie?

Aanpak

Deze vraag heb ik toegevoegd om zelf kennis te maken met alle metadata en om zelf te ervaren hoe het is om door alle beschikbare metadata te zoeken. Alle resultaten staan nog verder uitgebreid in de bijlage 'EDM in context'

Resultaat

Op dit moment wordt SharePoint gebruikt om het grootste deel van de metadata beschikbaar te maken voor gebruikers. Dat gaan om:

- Data domein en eigenaren lijst
- Business object + metadata
- Definities
- Datadomein kaart
- Informatie Product eigenaren en metadata
- (K)PI definities

Per beschikbare databron wordt uitgelegd hoe deze beschikbaar is gemaakt voor gebruikers:

Soort metadata	Beschikbaarheid
Data-domeinen en eigenaren lijst	Dit is een SharePoint lijst die te vinden is op de datamanagement pagina. Vanuit deze lijst is er geen koppeling naar andere databronnen. Het is mogelijk om de SharePoint lijst te exporteren als Excel bestand, het mist dan wel informatie over het datadomeincluster.
Business object + metadata	Dit is een SharePoint lijst die te vinden is op de datamanagement pagina. Een business object in de lijst kan een link bevatten naar de bijbehorende definitie. Deze link verwijst door naar een SharePoint pagina met de definitie die een link bevat naar de What is What. Ook is er een link naar de datadomein landkaart voor het weergeven van alle objecten in het datadomein van het business object.
Definities	De definities staan in het programma 'What is What'. Maar er is ook een SharePoint lijst die te vinden is op de datamanagement pagina. Een business object in de lijst kan een link bevatten naar de bijbehorende definitie. Deze link verwijst door naar een SharePoint pagina met de definitie.
Datadomein landkaart	De datadomein landkaart is ook te vinden via de datamanagement pagina. De landkaart heeft als doel om overzicht te geven over beschikbare metadata. Definities zijn hier niet in te vinden evenals de nieuwe soorten metadata die het

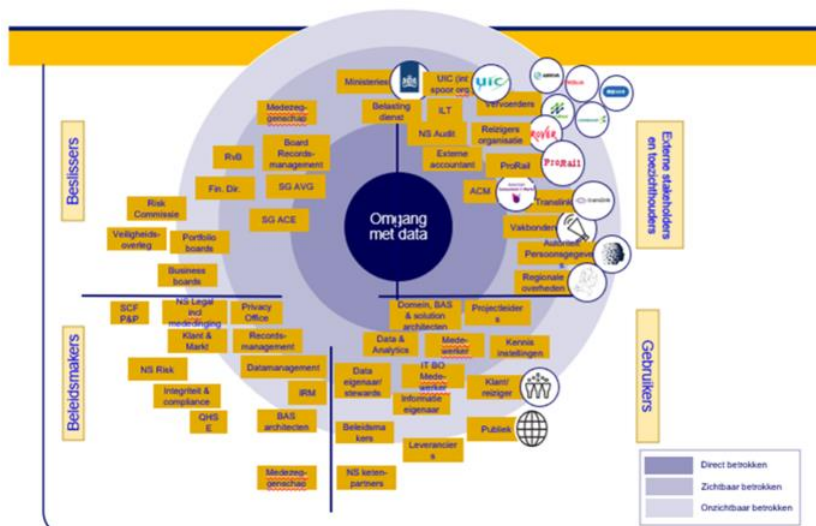
	EDM beheerd (IP). Er is op dit moment nog geen oplossing gemaakt die alle metadata in één overzicht weergeeft.
Informatie product en eigenaarschap	Het beheer van deze informatie is nieuw en daarom is de informatie nu nog niet te vinden op de datamanagement pagina. Deze informatie is te vinden via SharePoint en staat in een Excel bestand. In het Excel bestand staan ook business objecten die worden gebruikt maar er is geen directe link die verwijst naar de metadata lijst. Hetzelfde geldt voor KPI's die bij een IP horen.
(K)PI definitie	Deze informatie is te vinden op de datamanagement pagina. De KPI definities zijn opgesplitst per jaar. Elke definitie in een jaar heeft een eigen bestand waarin de definitie en eigenaren staan beschreven. Er is geen koppeling met andere metadata.

7.3 Wat zijn de belangrijkste gebruikersgroepen van datamanagement informatie?

Aanpak

Het originele plan was om een stakeholder analyse te doen. Dit was net nodig omdat het EDM-team al een stakeholder analyse heeft gedaan

Resultaat



Figuur 5 stakeholderanalyse

Het EDM-team heeft een stakeholder analyse ter beschikking, die wordt hierboven weergegeven. Deze analyse is te uitgebreid voor mijn afstudeerstage en daarom heb ik een aantal belangrijke groepen uit de analyse gehaald en verdeeld in categorieën. Deze categorieën zijn bepaald door de verschillende metadata types beschreven in (DAMA International, 2017):

- Business metadata
- Technische metadata
- Operationele metadata: Het EDM-team beheert geen operationele metadata dus deze wordt niet gebruikt.

Business metadata

Business metadata beschrijft data vanuit het business gebruik. Het beschrijft de business definities, regels en context voor data. Het zorgt voor een semantische laag tussen IT en business gebruikers.

Beschikbare business metadata bij het EDM team zijn:

- Data-domeinen en eigenaren
- Definities
- Business object en metadata
- (K)PI definities
- IP eigenaren en metadata
- Data scorecards en regels

Technische metadata

Technische metadata voorziet de gebruiker van technische details over data: wat is de structuur, format en regels voor het opslaan van data. Technische metadata helpt IT bij het beheren en groeien van systemen.

Beschikbare technische metadata bij het EDM-team zijn:

- Systeem informatie
- LDM
- Metadata over schema's, tabellen , attributen en relaties uit OPS, EDW en data marts
- ETL metadata zoals source-target mapping
- Data lineage
- Power BI metadata

Operationele metadata

Beschrijft details over de verwerking en toegang tot gegevens. Het kan bijvoorbeeld gaan over SLA requirements, error logs etc. Dit type metadata wordt niet verzameld door het EDM-team (behalve regels data gebruik)

Gebruikersgroepen

De gebruikersgroepen zijn opgesplitst in twee delen. De business gebruikers en de technische gebruikers. In de realiteit zal een technische gebruiker ook gebruik maken van business metadata en vice versa. In een organisatie zijn er een aantal rollen/groepen die gebruik maken van EDM metadata. Hieronder zijn deze weergegeven:

- Technisch
 - o Applicatieontwikkelaars
 - o Systeemanalisten
 - o Data warehouse ontwikkelaars
 - o BI ontwikkelaars
 - o Data integrators
 - o Data architecten
 - o Data scientists
- Business
 - o Business gebruikers
 - o Domein Architecten
 - o Kennis werkers
 - o Externe partijen
 - o Toezichthouders

7.4 Waardoor is het gebruik van EDM metadata minder dan de potentiële mogelijkheden?

Aanpak

Het doel was om mogelijke punten te ontdekken die kunnen worden verbeterd. Omdat ik voor deelvraag 2 al gebruik heb gemaakt van het metadata portaal kon ik met behulp van mijn stagebegeleiders een aantal belangrijke punten noemen die kunnen worden verbeterd.

Resultaat

Het huidige EDM portaal kan voor minder ervaren gebruikers als moeilijk toegankelijk worden ervaren mede door de hoeveelheid aan informatie. Hieronder een lijst met mogelijke problemen:

- Gebruikers hebben moeite met het vinden van informatie omdat ze niet precies weten wat ze nodig hebben.
- Er zijn geen duidelijke relaties tussen verschillende en dezelfde soorten metadata.
- Veel informatiebronnen missen en link naar andere bronnen waardoor een gebruiker zelf moet gaan zoeken naar de locatie van de andere bron.
- De beschikbare metadata is niet op één plek opgeslagen dit maakt het moeilijk voor een gebruiker om te navigeren door verschillende soorten metadata.
- Gebruikers moeten elke keer opnieuw zoeken naar belangrijke metadata die ze eerder hebben bekeken. Het is niet mogelijk om één keer iets te zoeken en het dan te bewaren.
- Iedereen ziet de informatie van de hele organisatie waardoor er een overvloed aan informatie is die niet relevant is voor elke gebruiker. De informatie past niet binnen de context van de individuele gebruiker.
- Er is geen startpagina waar een duidelijk overzicht wordt gegeven van alle mogelijke metadata.

Zodra het huidige traject succesvol geïmplementeerd is zijn er veel technische problemen opgelost en zijn er meer soorten metadata beschikbaar. Alle metadata staat dan opgeslagen in Axon, deze kan ook worden gebruikt als metadata portaal. Het nadeel van Axon dat dit een redelijk complexe omgeving is om te gebruiken en het is niet mogelijk om custom visualisaties te maken. Daarom zal SharePoint blijven worden gebruikt als metadata portaal.

Voor verdere uitleg zie Bijlage 'EDM in context'

7.5 Wat zijn wensen van medewerkers over het gebruik en toegankelijkheid van metadata?

Het originele plan was om wensen en eisen te verzamelen van medewerkers over hoe zij metadata gebruiken en hoe het metadata portaal daarbij kan helpen. Deze deelvraag is echter weggefallen

Gedeeltelijk omdat binnen NS weinig mensen veel kennis hebben van metadata. De mensen die deze kennis wel hebben zijn de data teams/ stewards. Deze hebben al contact met het EDM-team en de product owners en daarom zijn de eisen en wensen al bekend.

Ook kon ik ervan uitgaan dat de axon implementatie al was voltooid zoals vermeld in 'Aannames'. Alleen niemand heeft Axon nog gebruikt. Dit betekent dat ik het met gebruikers moet hebben over het metadata portaal in SharePoint terwijl een groot gedeelte van de problemen al zijn opgelost door Axon

Door deze redenen heb ik ervoor gekozen om niet direct met medewerkers wensen en eisen te verzamelen. Ik kom hier in mijn reflectie nog op terug.

7.6 Wat zijn de huidige best practices van metadatamanagement?

Aanpak

Om deze vraag te kunnen beantwoorden doe ik een literatuur/website onderzoek. In het plan van aanpak heb ik al vermeld dat er waarschijnlijk niet heel veel te vinden is over best practices van metadatamanagement.

Resultaat

Het was belangrijk om eerst een goed beeld te krijgen van metadatamanagement en daarna te zoeken naar best practices. Metadatamanagement en de best practices zijn uitgebreid beschreven in de externe bijlage 'metadatamanagement'.

Na het onderzoek zijn er een aantal best practices die door verschillende bronnen worden genoemd. Ik benoem ze in het kort. Deze staan ook in de externe bijlage 'metadatamanagement':

- Maak een datastrategie: Dit is essentieel om duidelijkheid voor jezelf en stakeholders te krijgen
- Gebruik metadata standaarden: Een metadata standaard helpt gebruikers omdat ze weten hoe ze van metadata gebruik kunnen maken
- Automatiseer zoveel mogelijk: Het is anders onmogelijk om organisatie breed metadata te verzamelen
- Democratiseer metadata: Metadata moet voor iedereen beschikbaar zijn. Als het voor iedereen beschikbaar moet zijn in de organisatie moet er rekening worden gehouden met de technische vaardigheden van medewerkers en het moet daarom makkelijk te gebruiken zijn
- Continue verbeteren van metadata: metadatamanagement is een organisatie breed proces en kan alleen werken als medewerkers hieraan bijdragen en EDM die bijdrage ook mogelijk maakt

- Focus op de metadata, niet op de tool: Met alleen een metadata tool kom je er niet. Een metadata repository vereist net als andere IT-initiatieven veel kennis en investeringen.

7.7 Hoe kan de bruikbaarheid (relevant, eenduidig en gebruiksvriendelijk) van het EDM metadata portaal worden gemeten?

Aanpak:

Origineel was de vraag 'Hoe kan de bruikbaarheid van EDM metadata worden gemeten'. Deze is vervangen met de huidige vraag aangezien de focus van het project lag op het portaal.

Resultaat

Ik heb nulmeting ontworpen en geen web usability test uitgevoerd omdat je bij deze test al een product moet hebben die gebruikers kunnen testen. Aangezien ik adviezen geef en niet voor elk advies een werkend eindproduct kan maken in mijn stageperiode is er daarom gekozen voor een nulmeting.

De nulmeting wordt gedaan over een metadata portaal. Dit betekent dat er geen vragen zijn over de opslag van data of over het verander proces. Het metadata portaal wordt gezien als system of engagement.

Ik heb weinig informatie kunnen vinden over het specifiek testen van een metadata portaal. De enige bron die ik heb kunnen vinden is (Marco, 2000). Hierin stonden een klein aantal vragen maar die zijn verouderd. Daarom heb ik gekeken naar meer standaard manieren om de bruikbaarheid van een website te testen. Hierdoor kan ik een nulmeting maken die gebaseerd is op de theorie van web usability en requirements van een metadata portaal.

De nulmeting wordt opgedeeld in 3 delen:

1. Gebruiksvriendelijkheid (maximaal 125 punten)
2. Eenduidig (maximaal 35 punten)
3. Relevantie (maximaal 35 punten)

Elke vraag in de meting hoort bij één van de bovenstaande categorieën. Elke vraag heeft een aantal mogelijkheden waarvan elke mogelijkheid x aantal punten (1-5) kan geven. Nadat alle vragen zijn beantwoord kan er per categorie een eindcijfer worden uitgerekend.

De nulmeting is te zien in de bijlage 'portaal test methode'.

8 Proof-of-concept

8.1 Functionaliteiten

In de bijlage 'advies' worden een aantal functionaliteiten gegeven voor het bruikbaar maken van het metadata portaal. Deze worden ook vermeld in de conclusie. Een van de deliverables is dat ik een proof-of-concept maak om een aantal van deze adviezen te verdedigen. Daarom heb ik met mijn begeleiders besproken welke functionaliteiten ik ga implementeren in het proof-of-concept.

Deze requirements worden neergezet in het formaat: "als gebruiker wil ik {functionaliteit} ". De rede dat ik dit formaat kies in plaats van de uitgebreidere versie waarbij de functie die de gebruiker heeft en de rede voor de functie is omdat er voor dit proof-of-concept maar één soort gebruiker is en de rede is al uitgebreid behandeld in de externe bijlage 'advies'.

Het prioriteren heb ik gedaan aan de hand van de MoSCoW-methode (European Knowledge Center for Information Technology, 2018)

Requirements: "als gebruiker wil ik {functionaliteit}"	Prioriteit (MoSCoW)
Als gebruiker wil ik kunnen abonneren en afmelden op metadata.	Must have
Als gebruiker wil ik mijn abonnementen kunnen inzien	Must have
Als gebruiker wil ik een groep kunnen aanmaken	Must have
Als gebruiker wil ik metadata aan een groep toevoegen en verwijderen	Must have
Als gebruiker wil ik lid kunnen worden van een groep	Must have
Als gebruiker wil ik metadata van een groep kunnen inzien	Must have
Als gebruiker wil ik een lijst van metadata kunnen zien (card)	Must have
Als gebruiker wil ik een lijst kunnen filteren via een zoekterm	Must have
Als gebruiker wil ik een gedetailleerd overzicht van een metadata object zien als ik op een item in de metadata lijst klik	Must have
Als gebruiker kan ik informatie krijgen over metadata	Must have
Als gebruiker wil ik een collectie aanmaken	Should have
Als gebruiker wil ik metadata toevoegen aan een collectie of verwijderen van een collectie	Should have
Als gebruiker wil ik alle metadata in een collectie inzien	Should have
Als gebruiker wil ik opmerkingen zien over een metadata object	Should have
Als gebruiker wil ik opmerkingen kunnen maken over een metadata object	Should have
Als gebruiker kan ik informatie krijgen over de verschillende metadata soorten	Should have
Als gebruiker wil ik in het gedetailleerde metadata overzicht kunnen doorklikken op gerelateerde metadata	Could have
Als gebruiker krijg ik extra informatie over attributen van een metadata object	Could have
Als gebruiker kan ik een rol kiezen	Won't have
Als gebruiker kan ik filters toepassen op metadata	Won't have
Als gebruiker kan ik metadata weergeven op verschillende manieren	Won't have
Als gebruiker kan ik extra functionaliteiten selecteren voor het beantwoorden van mijn use case	Won't have

8.2 Schets structuur

Hier beschrijf ik de architectuur voor het implementeren van de proof-of-concepts.

ETL-proces

In het plan van aanpak heb ik aangegeven dat ik een ETL-proces maak om data in te laden in de database. Maar omdat alle data die in de database gaat komen test data is die ik zelf heb gemaakt, is ervoor gekozen om geen ETL proces te maken en de data handmatig in de database te zetten.

Het zou dubbel werk zijn geweest om eerst zelf CSV of Excel bestanden te maken en vervolgens een ETL-proces te programmeren om de data in te laden. En de focus van dit project ligt niet aan de data integratie kant.

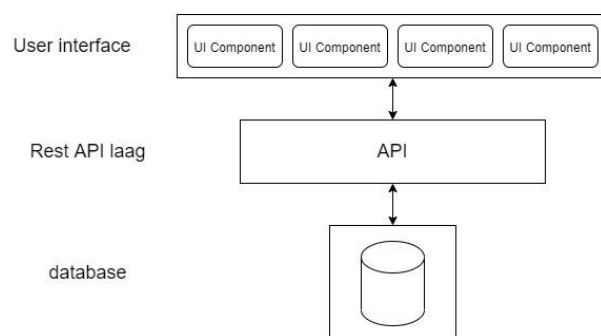
Overzicht software

Hiernaast is een vereenvoudigde weergave gemaakt van de architectuur.

Ik maak zelf test data aan de hand van een datamodel. Deze komen in een database (postgres) te staan

Nadat de database is gevuld met data wordt er een API laag omheen gebouwd die gebruikt kan worden om de metadata op een makkelijke manier op te halen en om extra functionaliteiten toe te voegen.

De laag daarboven “user interface” bevat het proof-of-concept. Die heb ik in de afbeelding UI componenten genoemd.



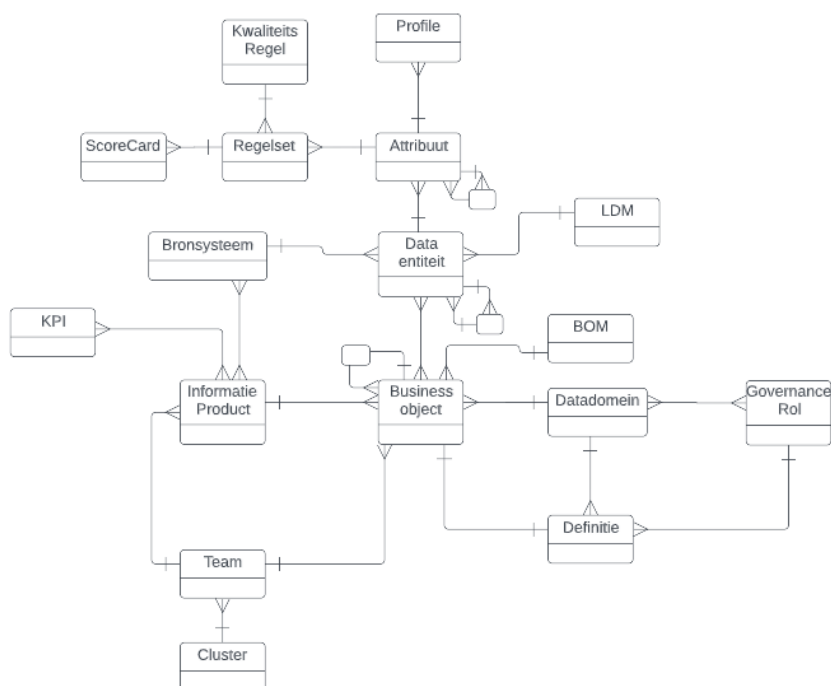
Figuur 6 Vereenvoudigde versie software architectuur

8.3 Data

In de Bijlage 'TDM' worden drie datamodellen beschreven:

- Huidig datamodel: Dit is hoe de data op dit moment opgeslagen is door het EDM-team.
- Conceptueel gegevens model: Dit is het model waar alle soorten metadata overzichtelijk met elkaar verbonden zijn.
- Technisch datamodel: Deze gebruik ik voor de database van het afstudeerproject.

Hieronder is het Conceptueel gegevens model te zien. Hierin is niet alleen de huidige beschikbare metadata maar ook de metadata na het traject naar Axon weergegeven. De modellen worden verder uitgelegd in de Bijlage 'TDM'.



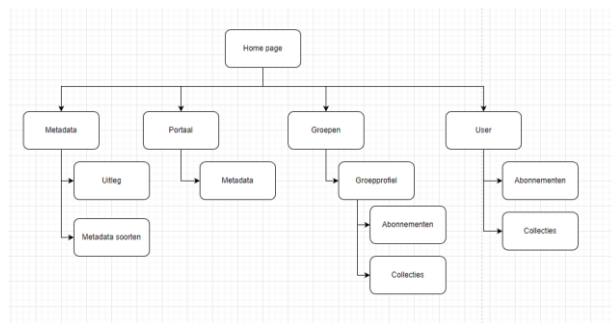
Figuur 7 CGM

8.4 Designs

Voor de verschillende requirements en de adviezen die worden gegeven in de externe bijlage 'advies' heb ik wireframes gemaakt om een beeld te geven aan de opdrachtgever hoe dit er in de praktijk uit zou kunnen zien. Ik heb voor wireframes gekozen omdat deze goed zijn in het weergeven van de functionaliteiten van een website, de structuur van een website en goed laten zien hoe informatie wordt weergegeven (usability, sd).

De designs zijn niet goed weer te geven in een word document en zijn daarom toegevoegd als externe bijlage 'designs'. Elk design bevat tekst ter verduidelijking.

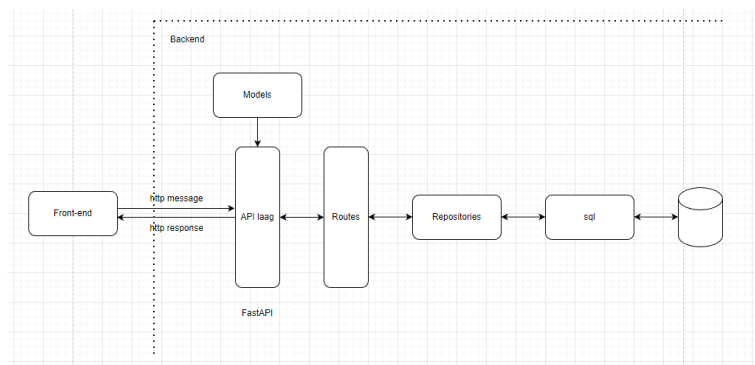
Ik heb ook een sitemap (Cardello, 2017) gemaakt voor het proof-of-concept om een blauwdruk van de website weer te geven.



Figuur 8 sitemap

8.5 API

Ik heb de API geprogrammeerd in python. Als web framework heb ik FastAPI gebruikt.



Figuur 9 API

Alle requirements genoemd in Functionaliteiten zijn geïmplementeerd behalve de requirements met prioriteit 'Won't have'. Op de volgende pagina zijn alle endpoints weergegeven van de API. Voor verdere uitleg over hoe de API werkt zie de externe bijlage 'technisch ontwerp'

FastAPI 0.1.0 OAS3
/openapi.json

Authentication ^

POST /auth/register Register

users ^

GET /users/ Read All Users

GET /users/{user_id} Get User

DELETE /users/{user_id} Delete User

GET /users/{user_id}/profiel Read User Profile

GET /users/{user_id}/profiel/collecties Get User Collections

GET /users/{user_id}/profiel/groepen Get User Groepen

groep ^

GET /groepen/ Get All Groups

POST /groepen/ Create Group

GET /groepen/{groep_id}/users Get Users In Group

GET /groepen/{groep_id}/collecties Get Collecties In Group

POST /groepen/{groep_id}/users/{user_id} Add User To Groep

DELETE /groepen/{groep_id}/users/{user_id} Remove User From Group

GET /groepen/{groep_id} Get Group

DELETE /groepen/{groep_id} Delete Group

collectie ^

POST /collectie/{abonnee_id}/ Create Collectie

GET /collectie/{abonnee_id}/{collectie_id} Read Collectie

DELETE /collectie/{abonnee_id}/{collectie_id} Delete Collectie

abonnement ^

GET /subscriptions/{abonnee_id}/ Get Subscriptions

POST /subscriptions/{abonnee_id}/ Create Subscription

DELETE /subscriptions/{abonnee_id}/ Delete Subscription

metadata ^

GET /metadata/ Get Metadata Objects

GET /metadata/{metadata_id} Get Detailed Metadata

GET /metadata/{metadata_id}/opmerkingen/ Get Opmerkingen

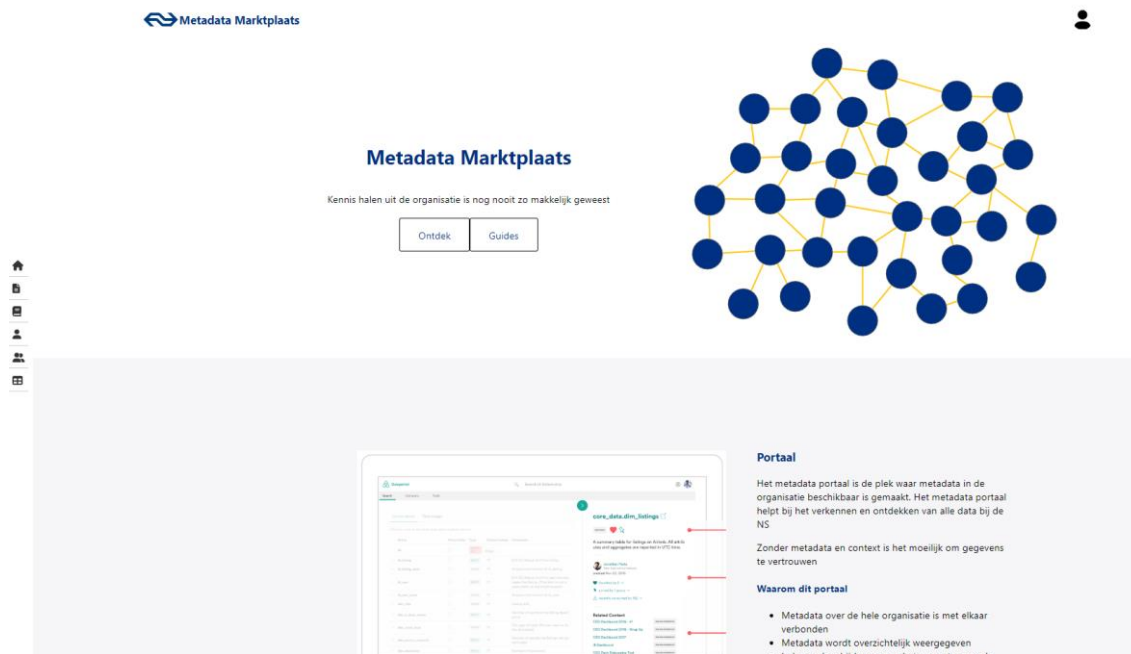
POST /metadata/{metadata_id}/opmerkingen/ Create Opmerking

POST /metadata/{metadata_id}/opmerkingen/{opmerking_id}/reactie Create Reactie

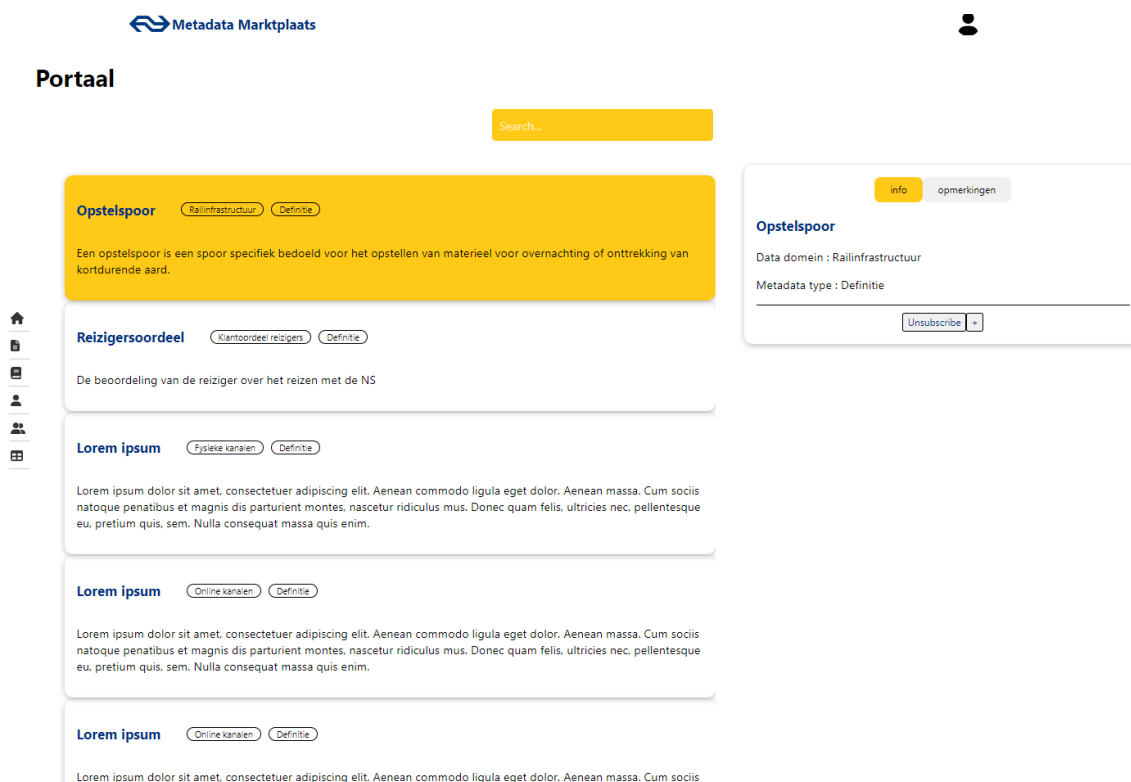
Figuur 10 endpoints van de API

8.6 Front-end

De website heb ik gemaakt met vue en typescript. Alle 'Must have' requirements zijn toegevoegd. Ook is de requirement 'Als gebruiker wil ik opmerkingen zien over een metadata object' toegevoegd. Ik heb aan mijn begeleiders een demo gegeven van de website. De website is moeilijk weer te geven op word maar de afbeeldingen van de website staan in de bijlage 'website'. Hieronder is de homepage te zien en het portaal waarin de metadata kan worden gevonden. Voor verdere uitleg over de front-end code zie de externe bijlage 'technische ontwerp'



Figuur 11 homepage website



Figuur 12 metadata portaal

9 Conclusie

“Hoe kan de NS EDM-afdeling de aanwezige EDM-informatie beschikbaar stellen voor inzage en gebruik binnen de NS,

zodanig dat deze relevant, eenduidig en gebruiksvriendelijk binnen de context van het gebruik ter beschikking is.”

De conclusie bestaat uit een aantal delen.

Relevant

Uit mijn onderzoek is gebleken dat het op dit moment niet mogelijk is om relevante metadata voor een gebruiker te bepalen aan de hand van de informatie die over gebruikers beschikbaar is.

Het is wel mogelijk om de gebruiker context te bieden door hem de mogelijkheid te geven om groepen/collecties aan te maken en te abonneren op metadata. Ook functionaliteiten zoals het opslaan en laden van filters kunnen de gebruiker helpen om zelf te bepalen wat relevant is. Dit vereist wel meer werk aan de kant van de gebruiker.

Eenduidig

Om de beschikbare EDM-informatie eenduidig te maken voor de gebruiker moet er meer metadata over de metadata worden weergegeven. Alle attributen van de beschikbare metadata soorten moeten worden duidelijk gemaakt om voor zo min mogelijk verwarring te zorgen.

Het portaal moet duidelijk maken in simpele termen wat metadata is en wat het nut is. Wat kan een gebruiker met metadata en wat zijn een aantal use cases waar metadata handig is. Zo krijgt de gebruiker meer inzicht en duidelijkheid.

Gebruiksvriendelijk

Het doel van een metadata portaal is niet om gebruikers zo lang mogelijk op de site te houden. Het doel is om te zorgen dat gebruikers zo snel en makkelijk mogelijk belangrijke metadata kunnen vinden en zo kennis kunnen vergaren. Daarom moeten er extra functionaliteiten worden ontwikkeld die hierbij helpen. Bekende use cases moeten al door het portaal worden afgehandeld en er moeten verschillende weergavemogelijkheden voor verschillende doeleinden worden toegevoegd om gebruikers te helpen.

Er zijn verschillende gebruikersgroepen die gebruik maken van metadata en er zit een groot verschil tussen onervaren en ervaren gebruikers. Het metadata portaal zou voor alle gebruikers moeten zijn maar zou ook extra functionaliteiten moeten hebben zodat ervaren gebruikers minder tijd hoeven te besteden naar het zoeken. Deze functionaliteiten zijn optioneel en worden niet automatisch weergegeven zodat nieuwe gebruikers niet worden afgeschrikt.

Als laatste moet zoveel mogelijk soorten metadata beschikbaar zijn. Als er relaties zijn tussen soorten metadata moeten die worden weergegeven en gebruikers moeten altijd van het ene metadata object naar het andere kunnen doorklikken. Dit verhoogt de kans dat gebruikers vinden wat ze nodig hebben.

10 Aanbevelingen

In de externe bijlage worden adviezen gegeven die de bruikbaarheid van metadata verhogen. Door gebruik te maken van de test methode kan je de verbetering zien. Ik heb het huidige metadata portaal (SharePoint) en een metadata portaal waar de adviezen in de bijlages zijn geïmplementeerd getest en zoals hieronder is te zien is het verschil in score groot. De scores zouden wat minder van elkaar verschillen als ik een test had gedaan na de Axon implementatie maar dat kan nog niet.

Huidig portaal:

- Gebruiksvriendelijkheid: 39/125
- Eenduidigheid: 20/35
- Relevantie: 5/35

Nieuw metadata portaal:

- Gebruiksvriendelijkheid: 110/125
- Eenduidigheid: 35/35
- Relevantie: 20/35

Mijn aanbeveling is niet om gelijk een nieuw metadata portaal te implementeren met de functionaliteiten in de bijlage 'advies'. Ik raad aan om te wachten op de Axon implementatie en dan feedback vragen aan de gebruikers.

Als duidelijk wordt dat er dingen ontbreken die gebruikers zouden helpen kan er worden nagedacht of een metadata portaal (system-of-engagement) kan worden ontwikkeld.

11 Discussie

In dit hoofdstuk worden een aantal mogelijk vervolgonderzoeken voorgelegd

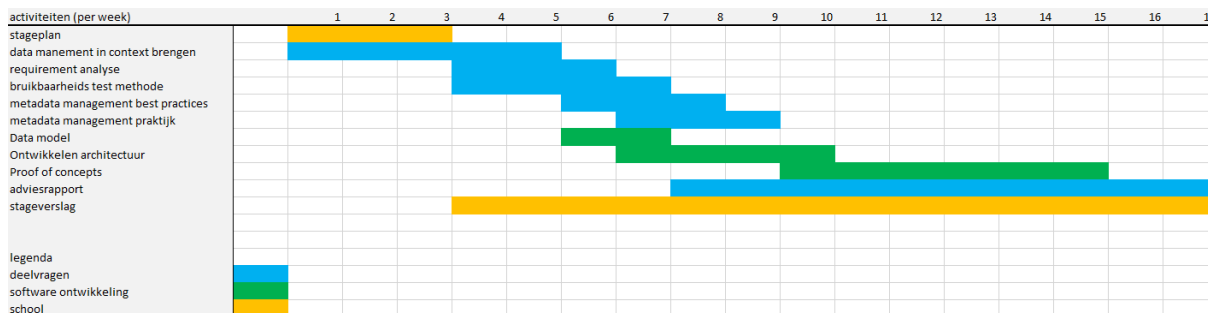
- Een vervolgonderzoek zou kunnen focussen op het krijgen van inzicht in hoeveel kennis medewerkers van de NS hebben over metadata. Zien ze het belang er wel van in.
- Onderzoek wat de use cases zijn voor metadata bij de verschillende gebruikersgroepen. Dit was origineel een gedeelte van mijn afstudeerstage maar is weggelaten.
- Een volgend verkennend onderzoek kan worden uitgevoerd naar het maken van een test methode voor een metadata portaal. De test methode die ik heb gemaakt kan nog worden uitgebreid en verbeterd.

12 Reflectie

In dit hoofdstuk reflecteer ik op mijn stageperiode bij het EDM-team.

12.1 Stageproces

De planning die ik heb gemaakt in de eerste weken van de stage was erg globaal. Ik had er al rekening mee gehouden dat niet alles zou gaan zoals ik zou verwachten. De planning liep wel ongeveer overeen met hoe de stage daadwerkelijk is verlopen.



Figuur 13 originele planning

Er zijn wel een aantal dingen die afwijken van de originele planning:

De test methode en metadatamanagement best practices zijn pas later in de stage beschreven. Dat kwam omdat ik die weken bezig was met het inventariseren van alle data en het data modelleren heeft meer tijd gekost dan verwacht. In mijn planning geef ik aan dat ik 2 weken ga data modelleren dit is in de praktijk rond de 4 weken geworden, inclusief data inventarisatie. Er waren meer soorten metadata dan ik had verwacht en het kostte tijd om al deze databronnen door te gaan.

Zoals is beschreven in Projectgrenzen was het niet mogelijk om metadata uit de Axon systemen te halen en deze structuur in te zien. Daarom moest ik zelf bedenken hoe de structuur eruit zou kunnen zien. Ook dit heeft extra tijd gekost.

Omdat het data modelleren langer duurden kon ik ook pas later beginnen aan het maken van de API. Ik heb daarom eigenlijk maar kort aan de front-end kant gewerkt. Ik heb hier wel van geleerd dat ik het datamodel veel kleiner had moeten houden en hem later uitbreiden als ik nog extra dingen wil laten zien. Dan had ik veel meer tijd gehad voor de front-end.

Ik heb ook gemerkt dat ik de hele stageperiode de wensen van NS heb moeten afwegen tegenover de wensen van school. Mijn opdrachtgever wilde veel activiteiten zoals; portaal designs en proof of concepts met zo min mogelijk werk aan de backend. Hiermee kan ik mijn B-competenties niet ontwikkelen, maar als ik me alleen richt op de B-competenties dan heeft de opdrachtgever uiteindelijk niks aan het resultaat.

12.2 Product

Ik heb niet al mijn doelen behaald tijdens de stage, maar ik ben alsnog tevreden over wat ik heb gedaan en geleerd. Tijdens de stage zijn er veel onderwerpen aan bod gekomen (modelleren, ui design, architectuur, backend, front-end en metadatamanagement) waardoor ik veel heb geleerd. Omdat ik uiteenlopende taken heb uitgevoerd had ik minder tijd voor de front-end en daarom heb ik de verduidelijking van mijn adviezen vooral gebaseerd op de designs. Ik denk dat de opdrachtgever een goed beeld heeft over hoe metadata bruikbaar kan worden gemaakt in de organisatie.

12.2.1 Requirements analyse

Zoals vermeld in Hoofdstuk 7.5 heb ik geen requirements analyse gedaan. Aan de ene kant denk ik dat ik hier de goede keuze heb gemaakt.

- Veel (niet-functionele) requirements kunnen indirect worden gehaald uit de mogelijke problemen in beschreven in 7.4.
- In mijn 'advies' wordt aangegeven dat het mogelijk moet zijn om verschillende use cases voor het zoeken van data toe te voegen aan het portaal.
- In de bijlage 'test methode' worden extra punten gegeven als de use cases die het portaal kunnen beantwoorden makkelijk kunnen worden uitgebreid als hier vraag naar is.

Voor een schets van het design is het dus niet nodig om naar specifieke use cases te kijken voor het zoeken naar metadata. Want het moet mogelijk zijn om use cases aan het portaal toe te voegen.

Het doen van een requirements analyse met verschillende medewerkers binnen NS had mij wel de mogelijkheid gegeven om kennis te maken met verschillende onderdelen van de organisatie. Achteraf gezien vind ik het jammer dat ik het niet heb gedaan.

12.3 Persoonlijke reflectie

Tijdens mijn afstudeerstage ben ik tegen een aantal persoonlijke punten gelopen waar ik nog aan moet werken.

Ik merk dat ik heel veel moeite heb met het voorspellen hoelang iets duurt om af te krijgen. Ik ga vaak van het meest positieve scenario uit. Hierin gaat niks mis of komen er onverwachte dingen bij. Ik heb regelmatig tegen mijn begeleiders gezegd hoelang een taak duurt en het duurde vrij vaak langer. Ik heb dit ook vermeld aan mijn begeleiders

Vooral aan het begin van mijn stage was ik constant bezig met het nadenken over alle mogelijkheden maar heb ik daar vervolgens niks mee gedaan. Na een tijd merkte ik dit over mezelf op en ook door de tijdsdruk ben ik toen veel meer gaan doen. Dit betekent niet dat denken verkeerd is maar het is wel handig om ook dingen te doen met de dingen die je hebt bedacht.

Ik heb tijdens mijn stage redelijk vaak gehad dat ik bezig was met een taak en vervolgens mij ging focussen op iets wat totaal niet belangrijk is. Vervolgens ben ik wel veel tijd aan dat niet belangrijke kwijtgeraakt. Het beste voorbeeld wat ik kan geven was toen ik de API en test data aan het maken was. Ik wou testcases schrijven voor het ophalen van metadata en het leek mij ineens een goed idee om code te schrijven voor het creëren van nep data voor de test database. Dit heeft mij een aantal dagen gekost voordat ik ermee was gestopt omdat het helemaal niet belangrijk was. Ik heb ervan geleerd dat ik vaker even stil moet staan en bedenken of dit wel de beste besteding van mijn tijd is.

13 Competenties

Per competentie leg ik uit hoe ik hieraan heb voldaan.

13.1 A-competenties

Onderzoek

- Vraagstelling: De hoofdvraag is SMART geformuleerd en de deelvragen behandelen de volledige breedte van de hoofdvraag.
- Methode en technieken: Per deelvraag leg ik de aanpak uit en onderbouw ik waarom ik voor deze methode heb gekozen.
- Informatieverzameling: Ik vermeld bronnen via de APA-wijze. De bronnen heb ik methodisch verzameld via de Big 6 methode uitgelegd in het Plan van aanpak.
- Analyse: Ik heb vermeld of ik kwalitatief/kwantitatief te werk ga en ik vermeld hoe ik tot de informatie van de deelvragen ben gekomen.
- Conclusies: De conclusie behandelt de breedte van de hoofdvraag en ik heb een hoofdstuk voor advies en vervolgonderzoek.

Leren leren

Ik heb individueel hulpbronnen verzameld door literatuuronderzoek voor het beantwoorden van deelvragen. Ook heb ik extra boeken gelezen en cursussen gedaan om mijn werkt te verbeteren. Deze vermeld ik hieronder:

- Boek: (Strengtholt, 2020)
- Cursus op udemy: Vue - The Complete Guide
- Cursus op udemy: Understanding TypeScript - 2022 Edition

Professioneel werken

- Probleemanalyse: Het probleem dat NS medewerkers te weinig gebruik maken van metadata is besproken.
- Planmatig werken: Ik heb een passende methode voor de aanpak van het gehele project gekozen en ook voor specifieke onderdelen is een aanpak bedacht.
- Methodisch werken: Ik heb methodisch gewerkt (zie Methodiek). Ook heb ik de planning actief bijgestuurd.

Reflectie: Ik heb constant feedback gekregen over de dingen die ik heb gemaakt. Ik heb ook gereflecteerd op persoonlijke valkuilen (zie Ik heb niet al mijn doelen behaald tijdens de stage, maar ik ben alsnog tevreden over wat ik heb gedaan en geleerd. Tijdens de stage zijn er veel onderwerpen aan bod gekomen (modelleren, ui design, architectuur, backend, front-end en metadatamanagement) waardoor ik veel heb geleerd. Omdat ik uiteenlopende taken heb uitgevoerd had ik minder tijd voor de front-end en daarom heb ik de verduidelijking van mijn adviezen vooral gebaseerd op de designs. Ik denk dat de opdrachtgever een goed beeld heeft over hoe metadata bruikbaar kan worden gemaakt in de organisatie.

13.1.1 Requirements analyse

Zoals vermeld in Hoofdstuk 7.5 heb ik geen requirements analyse gedaan. Aan de ene kant denk ik dat ik hier de goede keuze heb gemaakt.

- Veel (niet-functionele) requirements kunnen indirect worden gehaald uit de mogelijke problemen in beschreven in 7.4.

- In mijn 'advies' wordt aangegeven dat het mogelijk moet zijn om verschillende use cases voor het zoeken van data toe te voegen aan het portaal.
- In de bijlage 'test methode' worden extra punten gegeven als de use cases die het portaal kunnen beantwoorden makkelijk kunnen worden uitgebreid als hier vraag naar is.

Voor een schets van het design is het dus niet nodig om naar specifieke use cases te kijken voor het zoeken naar metadata. Want het moet mogelijk zijn om use cases aan het portaal toe te voegen.

Het doen van een requirements analyse met verschillende medewerkers binnen NS had mij wel de mogelijkheid gegeven om kennis te maken met verschillende onderdelen van de organisatie.

Achteraf gezien vind ik het jammer dat ik het niet heb gedaan.

- Persoonlijke reflectie)

Innovatie

De NS wou dat ik met een out-of-the-box idee kom om het huidige probleem aan te pakken. Daarom blijf ik op de hoogte van huidige ontwikkelen omtrent metadatamanagement.

Ook ben ik zelf met nieuwe ideeën en oplossingen gekomen tijdens mijn stage. De voor- en nadelen van deze ideeën zijn ook beschreven.

13.2 B-competenties

Opdracht in context plaatsen

- Organisatiekunde: Ik heb de organisatie beschreven, context gegeven en een aantal gerelateerde processen gemodelleerd zie hiervoor de externe bijlage 'EDM in context'.
- Data architectuur: Ik heb uitgebreide bijlages met informatie over de huidige data architectuur en die na het axon traject. Ook heb ik data modellen gemaakt om relaties terug te vinden.
- Datamanagement: Ik heb de functionele gebieden van data management beschreven en aangegeven welke van toepassing is op mijn opdracht.
- Eisen en wensen: Ik heb de eisen en wensen voor het proof of concept afgestemd.

Analyseren data architectuur:

- Data architectuur: Ik heb het huidige data architectuur gemodelleerd en beschreven. Ik heb de knelpunten aangegeven en ik heb een oplossing bedacht en gemodelleerd.
- Datamodelleren: Ik heb huidige data gemodelleerd en ik heb knelpunten aangegeven en vervolgens verbeteringen bedacht en toegepast.
- Data management: Ik heb de functionele gebieden van data management beschreven en aangegeven welke van toepassing is op mijn opdracht. Ook heb ik een meetinstrument ontwikkelt om op herhaalbare en overdraagbare wijze de huidige situatie van het metadata portaal te analyseren.
- Eisen en wensen: In de test methode komen de eisen en wensen omtrent een metadata portaal naar voren.

Ontwikkelen dataoplossing

- Data architectuur: architectuur van de oplossing is visueel weergegeven met een onderbouwing
- Data modeleren: Ik heb een datamodel gemaakt voor de proof of concept. Die past bij de requirements. Ik heb zelf standaarden bedacht en ik onderbouw ontwerpbeslissingen.
- Data storage: Ik heb een werkend prototype van de database ontworpen en gerealiseerd. Ik toen aan dat de database voldoet door test scripts te schrijven en een werkend proof-of-concept te laten zien.
- ETL-proces: zoals beschreven in hoofdstuk (Schets structuur) heb ik geen etl proces gemaakt. Dit was niet toepasselijk voor mijn stage opdracht
- Antwoorden: Ik maak een proof-of-concept dat metadata vragen kan beantwoorden

Adviseren dataoplossing

Ik geef advies over het gebruik van EDM-informatie. Mijn advies is gedeeltelijk gebaseerd op de best practices van metadatamanagement

- Context: Het advies is geplaatst in de context van de organisatie
- Advies: Ik kom met verschillende oplossingen voor het probleem. Geef mijn voorkeur en stem deze af met stakeholders
- Onderbouwing: Ik heb het advies onderbouwt met interne bronnen en externe bronnen.
- Methoden en technieken: Ik geef advies over metadatamanagement gebaseerd op de huidige best practices.

14 Bibliografie

Cardello, J. (2017, 8 22). *the modern web design process creating sitemaps and wireframes*.

Opgehaald van webflow: <https://webflow.com/blog/the-modern-web-design-process-creating-sitemaps-and-wireframes>

DAMA International. (2017). *DAMA-DMBOK: Data Management Body of Knowledge*. New Jersey: Technics publications.

European Knowledge Center for Information Technology. (2018, 5 8). *moscow-methode*. Opgehaald van ictportal: <https://www.ictportal.nl/ict-lexicon/moscow-methode>

Lean Six Sigma Groep. (sd). *Wat is Kanban*. Opgehaald van leansixsigmagroep: <https://leansixsigmagroep.nl/lean-agile-en-six-sigma/wat-is-kanban/>

Marco, D. (2000). *Building and Managing the Metadata Repository: A Full Lifecycle Guide*. Wiley.

Strengtholt, P. (2020). Data Management at Scale. In P. Strengtholt, *Data Management at Scale: Best practices for Enterprise Architecture* (p. 2). Californië: O'Reilly.

usability. (sd). *Wireframing*. Opgehaald van usability: <https://www.usability.gov/how-to-and-tools/methods/wireframing.html#:~:text=Wireframes%20serve%20multiple%20purposes%20by,intended%20functionality%20in%20the%20interface>

Wikipedia. (2022). *Web usability*. Opgehaald van Wikipedia: https://en.wikipedia.org/wiki/Web_usability

15 Externe bijlages

1. Afstudeervoorstel
2. Afstudeerplan
3. Studenten intervisie documenten
4. Bijlages
 - a. Advies
 - b. EDM in context
 - c. Metadatamanagement
 - d. Mogelijke oplossingen
 - e. Portaal test methode
 - f. Technisch data model
 - g. Technisch ontwerp
 - h. Website
 - i. Code
 - j. Data modellen
 - k. Designs