

MASTER'S THESIS

Hoe kunnen sturingsconcepten zoals FAIR Data Management, Data Governance, Enterprise-Architectuur en Community Models het management ondersteunen bij het sturen op hergebruik van publieke overheidsdata

Jonkman, Jeroen

Award date:
2023

[Link to publication](#)

General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain.
- You may freely distribute the URL identifying the publication in the public portal.

Take down policy

If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from <https://research.ou.nl/> on date: 28. Oct. 2023

Open Universiteit
www.ou.nl



Hoe kunnen sturingsconcepten zoals FAIR Data Management, Data Governance, Enterprise-Architectuur en Community Models het management ondersteunen bij het sturen op hergebruik van publieke overheidsdata

How can control concepts such as FAIR Data Management, Data Governance, Enterprise Architecture, and Community Models support management in guiding the reuse of public government data

Opleiding:	Open Universiteit, faculteit Betawetenschappen Masteropleiding Business Process Management & IT
Programme:	Open University of the Netherlands, faculty of Science Master of Science Business Process Management & IT
Cursus:	IM0602 Voorbereiden afstudeeropdracht BPMIT IM9806 Afstudeeropdracht Business Process Management and IT
Student:	ir. J.H.A. Jonkman
Identiteitsnummer:	
Datum:	28 februari 2023
Afstudeerbegeleider:	dr. ir. F.J.M. Mofers
Meelezer en examiner:	dr. L.H.H. Bollen
Versie:	1.0
Status:	Definitief

Abstract

De rol van data en sturing op data groeit binnen onze samenleving. Terwijl iedereen data kan produceren, delen en hergebruiken, is de manier waarop de data verzameld en opgeslagen wordt vaak niet traceerbaar. De Nederlandse overheid streeft naar het bevorderen van het hergebruik van publieke overheidsdata door middel van open data. Het Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) gebruikt de Data Agenda Overheid (DAO) om dit doel te verwezenlijken.

In dit onderzoek bekijken we in hoeverre de principes en richtlijnen van FAIR Data Management (FADAM), Data Governance (DG), Enterprise Architectuur (EA) en Community Models (CM) *Best Practices* kunnen bijdragen aan het verbeteren van de sturing op het beleid rond hergebruik van publieke overheidsdata. Door de combinatie van de principes en richtlijnen van FADAM, EA, DG is er een positieve impact op het verbeteren van de sturing op het hergebruik van publieke overheidsdata. Het is bekend dat de principes en richtlijnen van EA en DG het management ondersteunen bij sturen van bedrijfsprocessen rondom Big Data (BD). Dit onderzoek heeft aangetoond dat dit ook geldt voor publieke overheidsdata en dat FADAM hierin ook een belangrijke rol speelt in het verbeteren van de herbruikbaarheid. Bovendien hebben CM *Best Practices* een positief effect op gebruikersgroepen van publieke overheidsdata, waardoor het hergebruik verder wordt verbeterd.

Trefwoorden: FADAM, FAIR Data, FAIR Data Management, Data Governance, Enterprise Architectuur, Big Data, Overheid, Open Data, Community Models

Samenvatting

Data en de sturing op data spelen een steeds grotere rol binnen onze samenleving. Iedereen kan data produceren, delen en hergebruiken, maar de manier waarop de data verzameld, opgeslagen en eventueel hergebruikt wordt, is niet altijd traceerbaar en de herkomst, betrouwbaarheid en kwaliteit van de data is ook niet altijd bekend. Veel organisaties moeten omgaan met grote hoeveelheden data, bekend als Big Data (BD).

De Nederlandse overheid heeft een duidelijke taakstelling om het hergebruik publieke overheidsdata te bevorderen en deze zo toegankelijk mogelijk voor burgers, bedrijven en instellingen beschikbaar te stellen. Hiervoor worden publieke overheidsdata in de vorm van open data vrij beschikbaar gesteld en door verschillende groepen hergebruikt voor hun eigen bedrijfsprocessen. In dit onderzoek betekent publieke overheidsdata zowel uit open data als semi-open data. Het Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) geeft invulling aan de ambitie van de overheid om de mogelijkheden van data optimaal te benutten via de Data Agenda Overheid (DAO).

Om de problematiek beter te begrijpen is er een vooronderzoek uitgevoerd bij BZK om inzicht te krijgen in problemen die worden ondervonden bij het hergebruik van publieke overheidsdata binnen BZK. De vraag die naar voren kwam uit dit vooronderzoek naar voren kwam, kunnen de principes van FAIR Data Management (FADAM) aangevuld met algemene principes en richtlijnen van Enterprise Architectuur (EA) en Data Governance (DG) helpen om de sturing op het beleid rond hergebruik van publieke overheidsdata te verbeteren? Daarnaast kunnen de *Best Practices* van Community Models (CM) een positieve invloed hebben op het hergebruik van publieke overheidsdata. De hoofdvraag van dit onderzoek is:

Kan het toepassen van principes van FAIR Data Management aangevuld met algemene principes en richtlijnen, rond Enterprise Architectuur en Data Governance enerzijds en Community Models Best Practices anderzijds, helpen om de sturing op het beleid rond hergebruik van publieke overheidsdata te verbeteren?

Om deze vraag te beantwoorden zijn er zes literatuuronderzoeksvragen en drie empirisch onderzoeksvragen opgesteld. Als onderzoeksstrategie is gekozen voor een enkelvoudige casusstudie binnen de Nederlandse rijksoverheid. Na analyse van de gevonden literatuur zijn belangrijke kernwaarden geïdentificeerd die zijn gebruikt om het conceptueel model op te stellen. De onderzoeksmethode omvat interviews in twee verschillende afdelingen van de casus-organisatie: directie Digitale Overheid (DO) en Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP).

De resultaten van het onderzoek laten zien dat de principes en richtlijnen van FADAM kunnen bijdragen aan het verbeteren van beleidssturing rond het hergebruik van publieke overheidsdata. Bovendien is gebleken dat de algemene principes en richtlijnen van EA en DG invloed hebben op de sturing op beleid binnen de casus-organisatie. Het ontwikkelen van *Best Practices* van CM zal dit verder versterken. Voor de casus-organisatie wordt aanbevolen om zich in eerste instantie te richten op de metadata van publieke overheidsdata, waarbij principes en richtlijnen van FADAM en Master Data Management (MDM) een belangrijke rol speelt. Vanuit FADAM is *management commitment* een belangrijk aspect om het hergebruik van publieke overheidsdata te bevorderen.

Summary

Data and data management are playing an increasingly important role in society. Everyone can produce, share and reuse data, but the way the data is collected, stored, and potentially reused are not always traceable, and the origin, reliability and quality of the data is also not always known. Many organizations must deal with large amounts of data, known as Big Data (BD).

The Dutch government has a clear target to promote the reuse of public government data and to make this as accessible as possible of citizens, businesses, and institutions. To achieve this, public government data is made freely available in the form of open data and reused by various groups for their own business processes. In this research, public government data includes both open and semi-open data. The Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) implements the government's ambition to optimize the potential of data through the Data Agenda Overheid (DAO).

To better understand the issue, a preliminary research was conducted at BZK to gain insight into problems encountered in the reuse of public government data within BZK. The question that arose from this preliminary research was whether the principles of FAIR Data Management (FADAM) combined with general principles and guidelines of Enterprise Architectuur (EA) and Data Governance (DG) can help improve policy steering on the reuse of public government data. Additionally, the *Best Practices* of Community Models (CM) may have a positive impact on the reuse of public government data. The main research question of this research is:

Can the application of FAIR Data Management combined with general principles and guidelines of Enterprise Architectuur and Data Governance on one hand, and Community Models Best Practices on the other hand, help improve policy steering on the reuse of public government data?

To answer this question, six literature research questions and three empirical research questions have been formulated. A single case research within the Dutch government was chosen as the research strategy. After analyzing the found literature, important core values were identified and used to construct the conceptual model. The research method includes interviews in two different departments of the case organization: the directie Digitale Overheid (DO) and Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP).

The results of the research indicate that the principles and guidelines of FADAM can help improve policy steering on the reuse of public government data. Moreover, it was revealed that the general principles and guidelines of EA and DG have an impact on policy steering within the case organization. Developing best practices of CM will further strengthen this. For the case organization, it is recommended to focus initially on the metadata of public government data, where the principles and guidelines of FADAM and Master Data Management (MDM) play an important role. From FADAM, management commitment is an important aspect in promoting the reuse of public government data.

Inhoudsopgave

Abstract	ii
Samenvatting	iii
Summary	iv
1 Introductie	1
1.1 Achtergrond	1
1.2 Gebiedsverkenning	2
1.2.1 Enterprise Architectuur	2
1.2.2 Data Governance	3
1.2.3 Publieke overheidsdata	3
1.2.4 FAIR Data Management	3
1.2.5 Community Models	4
1.3 Probleemstelling	4
1.4 Opdrachtformulering	4
1.4.1 Onderzoeksvragen literatuuronderzoek	5
1.4.2 Onderzoeksvragen empirisch onderzoek	5
1.5 Motivatie / relevantie	6
1.6 Aanpak in hoofdlijnen	6
2 Theoretisch kader	8
2.1 Onderzoeksaanpak	8
2.1.1 Fase 1: Oriëntatie en voorbereiding	8
2.1.2 Fase 2: Literatuur verzamelen	9
2.1.3 Fase 3: Literatuur selecteren en beoordelen	9
2.1.4 Fase 4: Interpretieren en concluderen	9
2.2 Uitvoering.	9
2.2.1 Sturing middels Principes	10
2.3 Resultaten en conclusies	10
2.3.1 publieke overheidsdata	10
2.3.2 Enterprise Architectuur, Big Data en Open data	11
2.3.3 Data Governance, Big Data en Open data	13
2.3.4 FAIR Data Management, Big Data en Open data	14
2.3.5 Community Models, Big Data en Open data	15
2.3.6 Beantwoording literatuurvragen	15
2.3.7 Conceptueel model	16
2.4 Doel van het vervolgonderzoek	17
2.4.1 Vervolgonderzoek	18
3 Methodologie	19
3.1 Conceptueel ontwerp: keuze van onderzoeksmethode(n)	19
3.1.1 Onderzoeksmodel	19

3.2	Technisch ontwerp: uitwerking van de methode	20
3.2.1	Interviews	20
3.2.2	Bronnen	21
3.2.3	Plan van aanpak	21
3.3	Gegevensanalyse	21
3.4	Reflectie t.a.v. validiteit, betrouwbaarheid en ethische aspecten	22
3.4.1	Constructvaliditeit	22
3.4.2	Interne validiteit	22
3.4.3	Externe validiteit	22
3.4.4	Betrouwbaarheid	22
3.4.5	Ethische kwesties	23
4	Resultaten	24
4.1	Beschrijving onderzoeksproces.	24
4.1.1	Onderzoeksmethode	24
4.1.2	Introductie casus-organisatie	24
4.1.3	Uitvoering onderzoek	25
4.1.4	Selectiecriteria respondenten	26
4.1.5	Van theorie naar onderzoeksopzet	27
4.1.6	Aanpak data-analyse	27
4.2	Resultaten	27
4.2.1	Interview resultaten per variabele in het onderzoeksmodel	28
4.2.2	Resultaten documentanalyse	33
4.2.3	Samenvatting van de resultaten	34
5	Discussie, conclusies en aanbevelingen	36
5.1	Discussie - reflectie.	36
5.1.1	Toepassen van Enterprise Architectuur principes en richtlijnen	36
5.1.2	Toepassen van Data Governance principes en richtlijnen	37
5.1.3	Toepassen van FAIR Data Management principes en richtlijnen	37
5.1.4	Ontwikkelen van Best Practices Community Models	37
5.1.5	Sturing op beleid van publieke overheidsdata	37
5.1.6	Herbruikbaarheid van publieke overheidsdata verbeteren.	37
5.1.7	Reflectie onderzoeksproces	38
5.2	Conclusies	39
5.3	Aanbevelingen voor de praktijk	39
5.4	Aanbevelingen voor verder onderzoek	39
	Bibliografie	41
	Boek	41
	Academische Artikelen.	41
	Conferenties	42
	Technische Documentatie	43
	Acroniemen	44
	Woordenlijst	46
	Lijst van figuren	50
	Lijst van tabellen	51

A	Bijlage Verdieping	52
A.1	Gebiedsverkenning	52
A.1.1	Publieke overheidsdata	52
A.1.2	Hergebruikers	54
A.1.3	Motivatie / relevantie	55
A.2	Theoretische kader	56
A.2.1	Enterprise Architectuur Principes	56
A.2.2	Data Governance Principes	59
A.2.3	FAIR Data Management en FAIR Data Principes	61
A.2.4	Community Models	64
A.2.5	Beantwoording literatuurvragen	66
A.3	Methodologie	68
A.3.1	Conceptueel ontwerp	68
B	Bijlage Vooronderzoek	72
B.1	Introductie	72
B.1.1	Achtergrond	72
B.1.2	Gebiedsverkenning	73
B.1.3	Probleemstelling	74
B.1.4	Opdrachtformulering	74
B.1.5	Onderzoeksaanpak	74
B.2	Resultaat	74
B.2.1	Interview	75
B.2.2	Samenvatting interview	75
B.2.3	Analyse en evaluatie	76
B.3	Discussie en conclusie	76
B.3.1	Discussie	76
B.3.2	Conclusies	76
C	Bijlage Templates	77
C.1	Interview template	77
C.1.1	Samenvatting transcriptie interview	77
C.2	Template analyse documentatie	81
D	Bijlage Kwalitatieve interview data	82
D.1	Samenvatting transcripties	82
D.1.1	Enterprise Architectuur	82
D.1.2	Enterprise Architectuur en Big Data	83
D.1.3	Data Governance en Big Data	85
D.1.4	FAIR Data Management en Big Data	87
D.1.5	Community Models en Big Data	89
D.2	Resultaten triangulatie antwoorden	92
D.2.1	Analyse kernbegrippen en dimensies	92
D.3	Analyse documentatie	93
D.3.1	Wet openbaarheid van bestuur	93
D.3.2	Wet open overheid	94
D.3.3	Wet hergebruik van overheidsinformatie	95
D.3.4	Data Catalog Vocabulary	96

1

Introductie

1.1. Achtergrond

Data en de sturing op data spelen een steeds grotere rol binnen onze samenleving, iedereen kan data produceren, delen en hergebruiken. De herkomst, betrouwbaarheid en kwaliteit van data is niet altijd bekend. Ook de manier waarop data worden verzameld, vastgelegd en eventueel hergebruikt is niet altijd te herleiden. Voor een dataverwerkende organisatie was van belang dat data op een goed opgebouwde manier kunnen worden verwerkt door het opstellen van een zo goed mogelijk logisch en fysiek gegevensmodel. De meeste organisaties krijgen echter te maken met grote hoeveelheden data die niet vooraf systematisch opgebouwd kunnen worden. Het is daarom niet altijd duidelijk op welke manier deze data een rol kunnen vervullen in huidige en toekomstige bedrijfsprocessen.

Big Data (BD) is één van de termen waarmee deze ontwikkelingen aangeduid kan worden [Dedić, 2016]. Naast de ongestructureerdheid en de omvang van data, speelt de snelheid waarmee data beschikbaar komt een belangrijke rol. Daarnaast ook de inherente onzekerheid over de kwaliteit en de mate waarin data van waarde kunnen zijn voor de bedrijfsprocessen [Burmeister, 2018; Dedić, 2016]. Data Management (DM) is de ontwikkeling en executie van architecturen, regelgeving, praktijk, en procedures om op een goede manier de volledige data management lifecycle van een organisatie te ondersteunen [NL, 2021]. De non-profit organisatie DAMA ondersteunt organisaties bij het bereiken van hun data management doelstelling. Hiervoor heeft Henderson e.a. [Henderson, 2017] richtlijnen vastgelegd in het Data Management Framework (DAMA-DMBOK2). Data Management Lifecycle (DML) wordt als onmisbaar gezien voor de sturing op een strategisch niveau van BD [dFigueiredo, 2019].

De Nederlandse overheid heeft een duidelijke taakstelling om hergebruik van publieke overheidsdata te bevorderen en deze op een zo toegankelijk mogelijke manier beschikbaar te stellen aan burgers, bedrijven en instellingen. Voor vrij toegankelijke publieke gegevens wordt dit gedaan in de vorm van open data. Open data zijn voor iedereen toegankelijk, maar de mate van toegankelijkheid kan verschillen. Hiervoor is een classificatiesysteem voorgesteld door Berners-Lee [BernersLee, 2012], waarbij één ster staat voor een lage toegankelijkheid en vijf sterren voor een hoge mate van toegankelijkheid.

In 2019 heeft het Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) namens het kabinet de Data Agenda Overheid (DAO) gepubliceerd [Digitale Overheid, 2020a]. De DAO is een interbestuurlijke agenda, waarin staat wat de overheid gaat doen om beter om te gaan met persoonsgegevens, open data en BD.

Door publieke overheidsdata beter vindbaar te maken kan het hergebruik worden gestimuleerd voor de verschillende groepen hergebruikers welke elk hun specifieke wensen hebben [Donker, 2019] Binnen het domein van wetenschappelijke data zijn hiervoor enkele uitgangspunten beschreven zoals de FAIR Data Principles van Wilkinson e.a. [Wilkinson, 2016]. Door het toepassen van de FAIR Data Principles kan wetenschappelijke data

FAIR worden gemaakt. FAIR vertaald staat voor; vindbaar, toegankelijk, interoperabel en herbruikbaar [knav, 2018; Mons, 2011; Wilkinson, 2016].

Het hergebruik van data worden gestimuleerd door het effectief inzetten van Community Models (CM). Dit kan worden gedaan door vroegtijdige betrokkenheid van de belangrijkste betrokkene, een effectief bestuursmodel. Deze worden als belangrijke sociale aspecten gezien voor gemeenschapsontwikkeling [Curry, 2010].

De uitdaging waar BZK voor staat, kan als volgt worden samengevat: hoe kan de vindbaarheid en daarmee de herbruikbaarheid van publieke overheidsdata verder ondersteunen? Zijn de FAIR Data Principles hiervoor ook bruikbaar en kan bovendien het efficiënt inzetten van CM het hergebruik efficiënter maken [Curry, 2010; dFigueiredo, 2019]?

1.2. Gebiedsverkenning

In dit onderzoek staan de sturing op het beleid van publieke overheidsdata om het hergebruik te verbeteren centraal. De vraag hierbij is, kunnen de algemene principes en richtlijnen van Enterprise Architectuur (EA), Data Governance (DG) FAIR Data Management (FADAM) en kunnen de *Best Practices* van CM hierbij helpen?

Binnen dit onderzoek zijn enkele wetenschappelijke gebieden betrokken, deze zijn schematisch weergegeven in figuur '1.1-Gebiedsverkenning'.

In de volgende vier paragrafen komen eerste de algemene principes en richtlijnen van EA en DG aan bod. Daarna komen de aspecten van publieke overheidsdata aan bod. Ten derde komen de algemene principes en richtlijnen van FADAM aan bod en als laatste komen de *Best Practices* van CM aan bod.



Figuur 1.1: Gebiedsverkenning

1.2.1. Enterprise Architectuur

EA bestaat uit een samenvoeging van twee losstaande begrippen, Enterprise en Architectuur. Deze hebben hun eigen definities, zo beschrijft Architectuur de fundamentele concepten/eigenschappen van een systeem binnen zijn omgeving. Met Enterprise wordt een verzameling organisaties aangegeven die een gemeenschappelijk doel nastreven. EA verbindt deze twee begrippen om daarmee de missie, doelen en doelstellingen van de organisatie te ondersteunen. Dit wordt gedaan door het beschrijven van de processen en technische middelen die benodigd zijn om deze missie uit te voeren. Daarmee kan EA een totaalbeeld geven van een organisatie, maar als dit ontbreekt, wordt het lastig om als organisatie successen te behalen [Lankhorst, 2009].

Een van de bekendste methoden voor het ontwikkelen en beheren van de EA is de open standaard The Open Group Architecture Framework (TOGAF). Binnen de overheid wordt dit vastgelegd binnen de Nederlandse Overheid Referentie Architectuur (NORA). De principebeschrijvingen zijn gebaseerd op het format dat bekend is van TOGAF.

1.2.2. Data Governance

DG vormt de wetenschappelijke basis voor een redelijk jong onderzoeksveld binnen het Business Process Management (BPM) en maakt onderscheid tussen bestuur en management. Bestuur (governance) verwijst naar welke beslissingen genomen moeten worden om effectief beheer en gebruik van Informatietechnologie (IT) te verzekeren, zoals beslissingsdoelen. Daarnaast verwijst governance naar wie de beslissingen neemt zoals de plaats van verantwoording voor besluitvorming. Management omvat het nemen en uitvoeren van beslissingen [Khatri, 2010, p. 152]. De samenvoeging van data en governance omvat niet alleen het effectief beheer en gebruik van IT. Net zo belangrijk is het effectief inzetten van BD ontwikkelingen en daarmee de toegevoegde waarde van data voor de organisatie te behouden en/of te verbeteren.

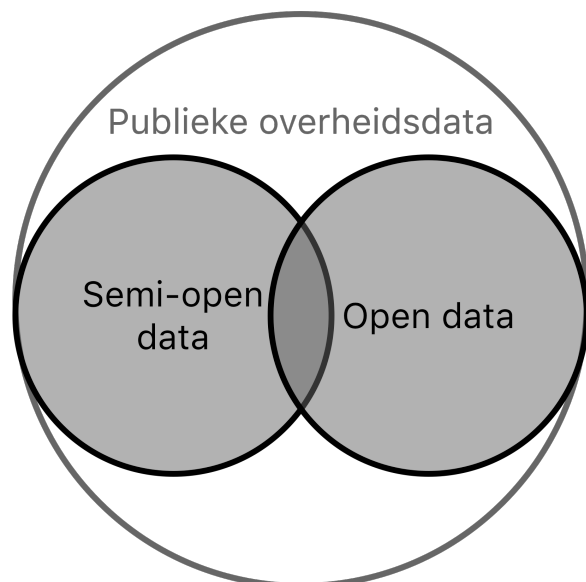
Binnen organisaties is DG steeds belangrijker geworden. Dit komt doordat er meer data worden gemaakt en verwerkt. De behoeften op het gebied van gegevensbeheer en data worden steeds meer gezien als essentiële bedrijfsmiddelen. De DAMA-DMBOK voorziet in behoeften die organisaties hebben voor het opzetten van DG. Binnen DAMA-DMBOK wordt het DAMA-DMBOK2 gebruikt als raamwerk voor DG [Henderson, 2017, Hoofdstuk 3.1].

1.2.3. Publieke overheidsdata

De publieke overheidsdata zijn onder te verdelen in open data en data die niet aan alle uitgangspunten voldoen van open data, ook wel semi-open data genoemd. Er is een eenduidige definitie te geven van open data. Binnen dit onderzoek worden de definities van publieke overheidsdata gebruikt, zoals deze zijn vastgelegd in bijlage 'A-Verdieping', paragraaf A.1.1.

Voor semi-open data is geen duidelijke definitie en er is geen volwassenheidsmodel beschikbaar. Binnen dit onderzoek wordt gebruik gemaakt van een eigen definitie van semi-open data zoals deze in bijlage 'A-Verdieping' paragraaf A.1.1 is uitgewerkt.

Standaarden die nu voor open data gelden waaronder Linked Data (LD) kunnen ook worden toegepast voor semi-open data. Big Open Linked Data (BOLD) is een samenvoeging van open data, BD en LD. Dit is een recent en snel opkomend werkveld volgens Janssen en Kuk [Janssen, 2016] en Janssen e.a. [Janssen, 2015]. Recentelijk is er door Lnenicka en Komarkova [Lnenicka, 2019] een aangepaste classificatie van BOLD gedefinieerd.



Figuur 1.2: Publieke overheidsdata

1.2.4. FAIR Data Management

FAIR is een afkorting van: vindbaar (Findable), toegankelijk (Accessible), interoperabel (Interoperable) en herbruikbaar (Re-usable). De FAIR Data Principles zijn ontworpen met een, datagedreven en door een machine ondersteunde open science architectuur als uitgangspunten. Met als doel om als leidraad te dienen voor DG voor onderzoekers en uitgevers in Academia [dFigueiredo, 2019, p. 14].

Recent wetenschappelijk onderzoek op het gebied van open overheid geeft aan dat FAIR Data Principles mogelijk toepasbaar zijn als raamwerk voor het DG binnen de publieke sector [dFigueiredo, 2019]. Hiervoor is het FADAM raamwerk een doorontwikkeling van de Figueiredo e.a. [dFigueiredo, 2019] voor de open overheid binnen de Verenigde Staten in combinatie met de zeven uitgangspunten van Tauberer en Lessig [Tauberer, 2020]. Daarmee zou het FADAM raamwerk een organisatie kunnen ondersteunen bij de sturing op het strategische niveau, maar ook bij de sturing op een meer operationeel niveau.

1.2.5. Community Models

Er zijn verschillende CM die kunnen worden toegepast om te communiceren met groepen stakeholders binnen een bepaalde datacommunities, in meer of mindere mate [Curry, 2010; KOOP, 2021a]. Zoals Curry e.a. [Curry, 2010] aangeeft, kan het volgen van deze *Best Practices* het gebruik van data worden vergroot. Dit zou ook kunnen gelden voor de open data communities om daarmee het hergebruik van publieke overheidsdata breder en efficiënter te maken.

1.3. Probleemstelling

De Nederlandse rijksoverheid en de decentrale overheid voeren gezamenlijk de DAO uit. Met DAO geeft het BZK invulling aan de ambitie van de overheid: de mogelijkheden van data zo goed mogelijk te benutten met oog voor publieke waarden [Digitale Overheid, 2020a]. Een van de uitdagingen op de agenda van de DAO, is het om overheidsdata kwalitatief te verbeteren en de herbruikbaarheid daarvan efficiënter te maken [Digitale Overheid, 2020a]. Dit kan worden bereikt door het inzetten van nieuwe technologieën op het gebied van data en ethiek [Digitale Overheid, 2020b].

Het doel van BZK is om het hergebruiken van publieke overheidsdata (open data) die de overheid beschikbaar stelt te laten hergebruik niet alleen door technici maar door iedereen. Nu worden open data met name gebruikt door data experts, maar hoe zorgen we dat we open data zo beschikbaar stellen dat deze door een veel bredere groep kan worden gebruikt.

Om de probleemstelling duidelijker te krijgen is er een vooronderzoek gehouden om inzicht te krijgen welke problemen worden ondervonden met het hergebruik van publieke overheidsdata binnen BZK. Voor het vooronderzoek is dezelfde sjabloon gebruikt als voor dit onderzoek. In het vooronderzoek wordt specifiek gesproken over open data, maar misschien gelden deze ook wel voor semi-open data en daarvoor ook voor publieke overheidsdata. Daarom is gekozen om de probleemstelling algemener te houden en publieke overheidsdata te gebruiken in plaats van open data en semi-open data. Het volledige vooronderzoek is in bijlage 'B-Vooronderzoek' opgenomen.

Het resultaat van dit vooronderzoek heeft geleid tot de een duidelijke vraagstelling. Kort samengevat laat de vraagstelling van dit afstudeeronderzoek zich omschrijven als volgt:

Er is een duidelijke behoefte aan sturing op het beleid van publieke overheidsdata om het hergebruik te verbeteren. Het is onbekend of de FAIR Data Principles hierbij kunnen helpen door publieke overheidsdata te verrijken met extra informatie. Welke data- en CM kunnen worden gebruikt om het hergebruik van publieke overheidsdata breder en efficiënter te maken.

1.4. Opdrachtformulering

Tijdens het vooronderzoek is een semi-gestructureerd interview gehouden met twee beleidsmakers van de afdeling directie Digitale Overheid (DO) die verantwoordelijk zijn voor de DAO. In dit interview is aangegeven dat Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) een goed startpunt is voor dit onderzoek. Binnen BZK is de afdeling DO verantwoordelijk voor het maken van beleid rond de DAO. De DAO heeft de verantwoordelijkheid om het hergebruik van open data te verbeteren en kan worden gezien als de probleemeigenaar. Onder andere door de aanbevelingen van Donker e.a. [Donker, 2019] is DO de opdrachtgever van dit onderzoek.

Binnen BZK zorgt KOOP namelijk voor de publicatie van alle officiële informatie van de Nederlandse overheid, waaronder de publieke overheidsdata [KOOP, 2021b]. KOOP ontwikkelt de standaarden om overheidsinformatie vindbaar te maken en werkt aan een transparante overheid door publieke overheidsdata beschikbaar te maken [KOOP, 2020]. Ook zorgt KOOP voor het contact met de verschillende gebruikersgroepen van publieke overheidsdata en voert het beleid uit van de DAO en is daarmee van essentieel belang voor dit onderzoek. Hierdoor kunnen de belangen van de hergebruikers van publieke overheidsdata, zie paragraaf A.1.2, toch worden meegenomen in dit onderzoek.

In dit afstudeeronderzoek staat de sturing op het beleid van publieke overheidsdata om het hergebruik te verbeteren centraal. Dit onderzoek heeft als doel te achterhalen welke aspecten van FADAM, DG en EA kunnen helpen met een goede sturing van FADAM, DG en EA, maar ook te achterhalen welke CM *Best Practices* kunnen er worden toegepast, met als doelstelling om daarmee het hergebruik van publieke overheidsdata te verbeteren.

Voor het literatuuronderzoek zal publieke overheidsdata te specifiek zijn en gebruiken we meer de algemene term van BD. Het onderzoek zal op een hoger abstractieniveau worden uitgevoerd, omdat het over de sturing op het beleid gaat. In de probleemstelling van paragraaf 1.3 die uit het vooronderzoek is gekomen, wordt gesproken over de FAIR Data Principles. Gezien dit onderzoek op een hoger abstractie niveau wordt uitgevoerd, is er voor gekozen om ook te kijken naar het FADAM raamwerk, waar de FAIR Data Principles onderdeel van zijn.

Dit onderzoek zal zich vooral toespitsen op de fase probleemanalyse en/of fase diagnose van de regulatieve cyclus. Daarmee valt het maken van een ontwerp dan wel het moduleren, implementeren en of evalueren van een (data) ontwerp buiten scope van dit onderzoek.

Daarmee is de hoofdvraag van dit onderzoek als volgt te formuleren:

Kan het toepassen van principes van FAIR Data Management aangevuld met algemene principes en richtlijnen, rond Enterprise Architectuur en Data Governance enerzijds en Community Models *Best Practices* anderzijds, helpen om de sturing op het beleid rond hergebruik van publieke overheidsdata te verbeteren?

1.4.1. Onderzoeksvragen literatuuronderzoek

Uit de bovenstaande hoofdvraag zijn in tabel '1.1-Onderzoeksvragen literatuuronderzoek' de onderzoeksvragen voor het literatuuronderzoek geformuleerd en voorzien van een L gevolgd door een volgnummer. Bij elke onderzoeksvraag is aangegeven in welke paragraaf het antwoord wordt behandeld.

Onderzoeksvragen voor het literatuuronderzoek		
L0.	Wat is het belang van het sturen middels EA/DG-principes?	§2.2.1
L1.	Welke principes en richtlijnen van EA kunnen helpen om de sturing op het beleid van BD het hergebruik te verbeteren?	§2.3.6
L2.	Welke principes en richtlijnen van DG kunnen helpen om de sturing op het beleid van BD het hergebruik te verbeteren?	§2.3.6
L3.	Op welke manier kan het FADAM raamwerk helpen het hergebruik van BD efficiënter te maken?	§2.3.6
L4.	Welke <i>Best Practices</i> van CM helpen het hergebruik van BD breder en efficiënter te maken?	§2.3.6
L5.	Op welke manier kan gestuurd worden op het beleid rond BD?	§2.3.6

Tabel 1.1: Onderzoeksvragen voor het literatuuronderzoek.

1.4.2. Onderzoeksvragen empirisch onderzoek

Uit de bovenstaande hoofdvraag en geformuleerde onderzoeksvragen voor de literatuuronderzoek zijn in tabel '1.2-Onderzoeksvragen empirisch onderzoek' de onderzoeksvragen voor het empirische onderzoek geformuleerd en voorzien van een E gevolgd door een volgnummer. Bij elke onderzoeksvraag is aangegeven in welke paragraaf het antwoord wordt behandeld.

Onderzoeksvragen voor het empirisch onderzoek		
E1.	Welke algemene principes en richtlijnen van EA en DG worden er gebruikt om te helpen om de sturing op het beleid rond het hergebruik van publieke overheidsdata te verbeteren?	§4.2
E2.	Welke principes van FADAM worden er gebruikt om te helpen om de sturing op het beleid rond het hergebruik van publieke overheidsdata te verbeteren?	§4.2
E3.	Welke <i>Best Practices</i> van CM worden er gebruikt om te helpen om het hergebruik van publieke overheidsdata te verbeteren?	§4.2

Tabel 1.2: Onderzoeksvragen voor het empirisch onderzoek.

1.5. Motivatie / relevantie

In een recentelijk adviesrapport van de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) is aangegeven dat de toegankelijkheid en herbruikbaarheid van publieke overheidsdata verbeterd kan worden door onder andere de FAIR Data Principles toe te passen [knaw, 2018]. Daarnaast zijn in het regeerakkoord van 2017–2021 plannen opgenomen over de open overheid. Deze plannen zijn verder vormgegeven door het DAO van de DO, door invulling te geven aan het verbeteren van het hergebruik van publieke overheidsdata [Digitale Overheid, 2020a].

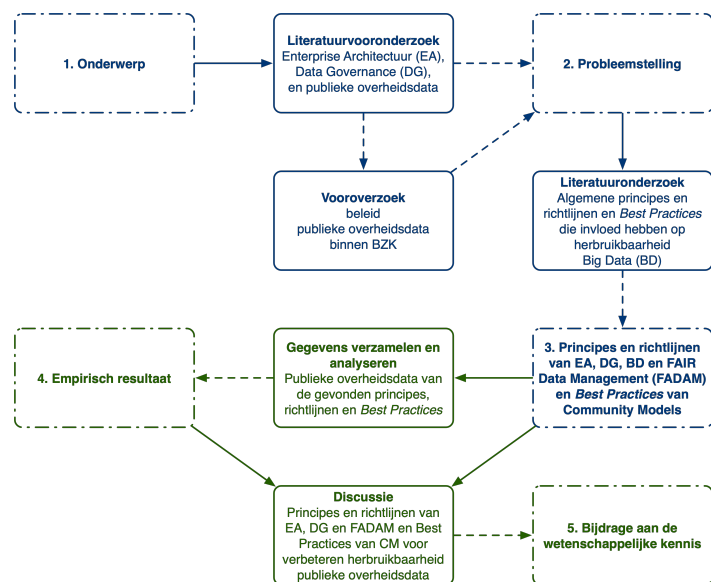
Voor het afstemmen van de FAIR Data Principles en de DAMA-DMBOK voor de open overheid binnen de Verenigde Staten is het FADAM raamwerk ontwikkeld door de Figueiredo e.a. [dFigueiredo, 2019]. Het FADAM raamwerk kan de toegankelijkheid en herbruikbaarheid van publieke overheidsdata mogelijk verbeteren door invulling te geven aan de EA/DG principes. Samenwerking en een goede communicatie met verschillende groepen die publieke overheidsdata hergebruiken kan de gebruiksvriendelijkheid en herbruikbaarheid van publieke overheidsdata positief beïnvloeden [Donker, 2019]. *Best Practices* van CM kunnen mogelijk ondersteunen bij het breder en efficiënter hergebruiken van publieke overheidsdata.

1.6. Aanpak in hoofdlijnen

Het onderzoek kent een gefaseerde aanpak volgens de Design Science Research Methode (DSRM) [Peppers, 2007].

Na het literatuurvooronderzoek is er een vooronderzoek gedaan om te bepalen of de gevonden aanbevelingen van recente onderzoeken ook worden herkend binnen BZK (bijlage B). Vervolgens is de probleemstelling geformuleerd waaruit de opdrachtformulering volgt om daarmee de hoofdvraag van dit onderzoek te formuleren.

De stappen van deze voorbereiding afstuderen Business Process Management & IT (BPM&IT) zijn hieronder kort benoemd en schematisch in het blauw weergegeven in figuur 1.3. De stappen van onderzoeksfase BPM&IT zijn schematisch in het groen weergegeven in figuur 1.3.



Figuur 1.3: Onderzoekstappen

In hoofdstuk 2 wordt het theoretisch kader uitgewerkt in een onderzoeksanpak, uitvoering en resultaten en conclusies. Waarna het doel van het vervolgonderzoek kan worden bepaald. In hoofdstuk 3 wordt de methodologie voor het vervolgonderzoek weergegeven in een conceptueel ontwerp een conceptueel model. Daarna wordt in het technisch ontwerp de methode uitgewerkt en wordt kort stil gestaan bij de gegevensanalyse. Als laatste de reflectie van validiteit, betrouwbaarheid en ethische aspecten. Voor het empirisch onderzoek van hoofdstuk 4 is gekozen om hiervoor semigestructureerde interviews af te nemen omdat er dan ruimte wordt gelaten voor verdieping [Saunders, 2019, p.436-444]. Het onderzoek wordt afgesloten met hoofdstuk 5 waarbij de discussie en conclusies worden behandeld die op basis van het empirisch onderzoek aan het licht zijn gekomen. Tot slot volgen er aanbevelingen voor verder onderzoek op dit onderwerp.

2

Theoretisch kader

In dit hoofdstuk wordt de aanpak beschreven van het literatuuronderzoek voor het ontwikkelen van het theoretische kader om de herbruikbaarheid van publieke overheidsdata te verbeteren. Het theoretisch kader richtte zich op het verkrijgen van inzicht in de algemene principes en richtlijnen van Enterprise Architectuur (EA) en Data Governance (DG), het toepassen van het FAIR Data Management (FADAM) raamwerk en *Best Practices* van Community Models (CM), met als doelstelling de sturing op het hergebruik van Big Data (BD) te verbeteren.

Organisaties in de publieke sector, zoals de rijksoverheid, kijken in bepaalde opzichten af van bedrijven met winstoogmerk. Zo is er geen sprake van het streven naar winst, maar wordt gestuurd op het toevoegen van publieke waarde voor de samenleving. Eén van deze waarden voor de samenleving is de herbruikbaarheid van publieke overheidsdata, waarbij publieke overheidsdata ook zeker gezien kan worden als BD. Dit doet de rijksoverheid niet alleen, maar met haar partners en met andere betrokkenen in de samenleving.

In de literatuur komen voornamelijk alleen de termen open data en BD voor. Binnen dit onderzoek gebruiken we de term publieke overheidsdata omdat het meer dan alleen open data betreft, zoals ook semi-open data. Als we specifiek BD of open data bedoelen, dan gebruiken we deze termen, anders gebruiken we de term publieke overheidsdata.

2.1. Onderzoeksaanpak

Het theoretisch kader geeft inzicht in eerder uitgevoerd onderzoek naar EA en DG algemene principes en richtlijnen en het toepassen van FADAM principes enerzijds en anderzijds het toepassen van *Best Practices* van CM. Centraal staat het beantwoorden van de literatuuronderzoeksvragen over het verbeteren van het hergebruik van publieke overheidsdata.

De aanpak van het literatuuronderzoek is verdeeld in vier fasen zoals deze zijn beschreven in Saunders e.a. [Saunders, 2019]. De theorie wordt verkregen met behulp van: Google Scholar (GS) en Digitale Bibliotheek Open Universiteit (DBOU).

2.1.1. Fase 1: Oriëntatie en voorbereiding

In de eerste fase zijn er een aantal publicaties geanalyseerd over eerder uitgevoerd onderzoek op het gebied van EA, DG, FADAM, CM in combinatie met BD en publieke overheidsdata. Deze zijn gebruikt in de gebiedsverkenning van paragraaf '1.2-Gebiedsverkenning'.

2.1.2. Fase 2: Literatuur verzamelen

Op basis van deze eerst geselecteerde publicaties en de onderzoeksvragen van tabel '1.1-Onderzoeksvragen literatuuronderzoek' zijn de volgende kernbegrippen geselecteerd:

- Big Data, voor het beantwoorden van L2 tot en met L5.
- Enterprise Architectuur, voor het beantwoorden van L1 en L5.
- Data Governance, voor het beantwoorden van L2 en L5.
- FAIR Data Management, voor het beantwoorden van L3.
- Community Models, voor het beantwoorden van L4.

2.1.3. Fase 3: Literatuur selecteren en beoordelen

In deze fase worden de gevonden artikelen beoordeeld op kwaliteit [Saunders, 2019]. Als selectiefilter zijn alleen artikelen van het afgelopen decennium genomen, dus die tussen 2010 en 2020 zijn gepubliceerd. De resultaten worden gesorteerd op relevantie, alleen de eerste 100 gevonden artikelen worden geanalyseerd. Daarnaast wordt gekeken naar het aantal references in combinatie met de publicatiedatum of een artikel ook wordt geselecteerd. De inhoud van het artikel moet te lezen zijn qua taal en het artikel moet vrij beschikbaar zijn in GS of DBOU.

- Eerst wordt er gezocht op de gedefinieerde zoektermen in combinatie met BD en open data.
- Van de artikelen wordt eerst alleen de samenvatting gelezen.
- In deze fase wordt de publicatie bestudeerd en beoordeeld op:
 - Compleetheid: hoe is het onderzoek uitgevoerd?
 - Eenzijdigheid: zijn er voldoende bronnen gebruikt?
 - Status auteur(s): wie is de auteur, wat heeft hij gepubliceerd?
 - Kwaliteit: zijn de argumenten goed onderbouwd, met behulp van bronvermelding?

2.1.4. Fase 4: Interpreteren en concluderen

Publicaties die gedurende de vorige fase zijn geselecteerd worden bestudeerd en geanalyseerd. Het programma Endnote is gebruikt voor het vastleggen van de references van de zoekresultaten. Endnote is verder gebruikt voor het verder beheer van alle references tijdens de literatuurstudie. Bovendien ondersteunt het programma de vastlegging van citaten en references volgens de APA-stijl.

2.2. Uitvoering

Op basis van de geselecteerde kernbegrippen van paragraaf '2.1.2-Fase 2: Literatuur verzamelen' zijn de zoektermen samengesteld zoals deze in tabel '2.1-Uitvoering literatuuronderzoek' zijn opgenomen.

Zoektermen	GS		DBOU	
	In artikel	In titel	In artikel	In titel
("enterprise architecture") AND ("big data")	6820	11	728	4
("enterprise architecture") AND ("open data")	1510	6	117	1
("data governance") AND ("big data")	11800	219	2182	31
("data governance") AND ("open data")	3270	32	392	0
("fair") AND ("big data")	17200	43	21894	19
("fair") AND ("open data")	17600	42	2823	5
("community model") AND ("big data")	875	1	140	0
("community model") AND ("open data")	264	0	51	0
("reuse") AND ("big data") ¹	56600	34	6658	15
("reuse") AND ("open data") ¹	32300	73	4123	18

Tabel 2.1: Uitvoering literatuuronderzoek

In de volgende vijf paragrafen wordt kort aangegeven waarom de geselecteerde literatuur belangrijk kan zijn voor het beantwoorden van de bijbehorende onderzoeksvraag of vragen.

2.2.1. Sturing middels Principles

In het DYnamic Architecture (DYA)-architectuurraamwerk worden er: (1) algemene principes, (2) beleidslijnen en (3) modellen, uitgezet tegen de bedrijfsdoelstellingen [vdBerg, 2004, p. 50]. Voor de algemene principes geldt dat deze de gezamenlijk visie van de business- en topmanagement weerspiegelen. Deze zouden zo geformuleerd moeten zijn dat deze door iedereen in de organisatie begrepen worden [OU, 2009, p. 140].

De algemene principes worden vertaald naar beleidslijnen die daarmee invulling geven aan de verschillende deelarchitecturen. Op dit niveau bevinden zich ook de standaarden en richtlijnen [OU, 2009, p. 140].

In de The Open Group Architecture Framework (TOGAF) standaarden wordt onderscheid gemaakt in *Enterprise Principles* en *Architecture Principles* [Josey, 2016, chap. 23]. Waarbij de enterprise principes de sturing op het beleid van de gehele organisatie betreft. De architectuurprincipes gaan vooral over een deel- of domeinarchitectuur.

De gewenste doelarchitectuur kan worden gevisualiseerd door de modellen [OU, 2009, p. 140]. Deze modellen bevinden zich vooral op de uitvoeringslaag en zijn daarom voor dit onderzoek minder van belang.

In het functionele raamwerk van DAMA-DMBOK worden op verschillende domeinen principes en richtlijnen gegeven die de organisatie kunnen help een goede DG in te richten [Henderson, 2017].

Hiermee kan de vraag van L0 worden beantwoord en worden gebruikt voor het verwerken van de resultaten van de literatuurstudie:

Wat is het belang van het sturen middels EA/DG-principes?

- Voor het vervolg van dit onderzoek en het literatuuronderzoek kijken we vooral naar de algemene principes en de vertaling hiervan naar beleidslijnen [Josey, 2016; OU, 2009; vd-Berg, 2004].
- Daarnaast kijken we ook naar de domeinen principes en richtlijnen van enkele specifieke architectuurprincipes die met de geselecteerde kernbegrippen van dit onderzoek te maken hebben [Henderson, 2017; Josey, 2016; OU, 2009].

In de uitvoering van het literatuuronderzoek wordt verder ingegaan op de volgende onderwerpen, om daarmee de overige literatuur onderzoeksvragen te beantwoorden:

- Enterprise Architectuur, Big Data en Open data
- Data Governance, Big Data en Open data
- FAIR Data Management, Big Data en Open data
- Community Models, Big Data en Open data

2.3. Resultaten en conclusies

In deze paragraaf worden de antwoorden gegeven op de literatuuronderzoeksvragen L.1 tot en met L.5 uit hoofdstuk '1-Introductie', tabel '1.1-Onderzoeksvragen literatuuronderzoek'. Als eerste staan we kort stil bij publieke overheidsdata. Daarna worden de onderwerpen geanalyseerd, zoals aangegeven in paragraaf '2.2-Uitvoering'. Deze paragraaf wordt afgesloten met het beantwoorden van de literatuuronderzoeksvragen.

2.3.1. publieke overheidsdata

Voor publieke overheidsdata is geen eenduidige term binnen de wetenschap. In dit onderzoek maken we een opdeling van publieke overheidsdata op basis van de toegankelijkheid in open data en semi-open data. In het artikel van Hasnain en Rebholz-Schuhmann [Hasnain, 2018], wordt er over Linked Open Data (LOD) van Berners-Lee [BernersLee, 2006] gesproken, deze komt beter overeen met de definitie van open data dan

¹Deze zoektermen zijn toegevoegd ter controle of er mogelijk belangrijke literatuur gemist is.

de vijf sterren stappen [BernersLee, 2012]. In recentelijk onderzoek door Lnenicka en Komarkova [Lnenicka, 2019] is er een aangepaste classificatie van open data gedefinieerd als Big Open Linked Data (BOLD).

Zoals eerder aangegeven, wordt in paragraaf 1.4 geen specifiek onderscheid gemaakt tussen BD en publieke overheidsdata enerzijds, en open data en publieke overheidsdata anderzijds in dit onderzoek. Het enige onderzoek dat een onderscheid maakt in open data en semi-open data is het onderzoek van Bargh e.a. [Bargh, 2016]. Hierin worden enkele aanbevelingen gedaan die we in dit onderzoek volgen, zoals het maken van een duidelijk onderscheid. In paragraaf 'A.1.1-Open data' is uitgelegd wat open data precies inhoudt en in paragraaf 'A.1.1-Semi-open data' is uitgelegd wat semi-open data precies inhoudt.

2.3.2. Enterprise Architectuur, Big Data en Open data

EA helpt bij het realiseren van strategieën van een organisatie. Een architectuur, in het regulatieve perspectief, wordt vormgegeven door principes en richtlijnen [Bommel, 2007]. Binnen dit onderzoek richten we voornamelijk op de algemene EA principes en de vertaling hier van in de EA richtlijnen.

De concepten van het modelleren van EA zijn beschreven in het onderzoek van Jonkers e.a. [Jonkers, 2004]. Hierbij wordt onderscheid gemaakt in drie verschillende concepten voor architectuurbeschrijvingen. Voor dit onderzoek zijn de business layer concepten belangrijk, omdat hierbij wordt gekeken naar bedrijfsprocessen, bedrijfsfuncties of zakelijke interacties [Jonkers, 2004, ch. 4].

In het onderzoek van Bommel e.a. [Bommel, 2007] worden vier fundamentele uitdagingen aangegeven voor het ontwikkelen van EA. Wat voor dit onderzoek mogelijk van belang kan zijn, de regulatieve behoeften en de formuleringsstrategie voor het bepalen van algemene EA principes.

Het onderzoek van Kotusev [Kotusev, 2017] beschrijft een conceptueel model van Enterprise Architecture Management (EAM), interessant hierin is hoe de besluitvormingsprocessen van een gedecentraliseerde organisatie worden beschreven. Dit model van EAM is misschien goed toepasbaar in dit onderzoek, omdat in dit model zowel de algemene EA principes als de domeinspecifieke principes en richtlijnen zijn opgenomen.

Het onderzoek van Gong en Janssen [Gong, 2021] is mogelijk interessant voor dit onderzoek, omdat dit een combinatie is van EA en BD. Omdat worden hierin enkele EA rollen en EA capabilities beschreven die van belang zijn met name voor Big Data Management (BDM).

Het onderzoek van Lnenicka en Komarkova [Lnenicka, 2019] zijn voorstellen gedaan voor een overheids-EA raamwerk om BOLD te ondersteunen met behulp van cloud computing. Mogelijk is dit EA-raamwerk goed toepasbaar in dit onderzoek voor de specifieke architectuurprincipes die relevant zijn voor BD.

Enterprise Architectuur referentiemodel

Het referentiemodel voor open gedistribueerde verwerking Reference Model for Open Distributed Processing (RM-ODP) is een gezamenlijke ISO/ITU-T standaard voor de specificatie van Open Distributed Processing (ODP). Het definieert vijf gezichtspunten op een ODPsysteem die elk hun eigen specificatietaal hebben. Voor het enterprise-gezichtspunt dat het doel, de reikwijdte en het beleid van een systeem beschrijft, is bijvoorbeeld de RM-ODP Enterprise Language gedefinieerd waarin bijvoorbeeld bedrijfsdoelstellingen en bedrijfsprocessen kunnen worden gemodelleerd [Jonkers, 2004].

Dit onderzoek richt zich op onder andere op architectuurprincipes en richtlijnen en niet op het modelleren of beschrijven van de EA. Daarom nemen we dit referentiemodel niet mee in het empirisch onderzoek.

Enterprise Architecture Management

EA is een beschrijving van een onderneming vanuit een geïntegreerd bedrijfsperspectief en Informatietechnologie (IT)-perspectief. EAM is een managementpraktijk die alle managementprocessen omvat die verband houden met EA en gericht is op het verbeteren van de afstemming tussen de business en IT [Kotusev, 2017]. Naast de traditionele vierstappenbenadering [Sowa, 1992; Zachman, 1987] zijn er nog twee EA raamwerken Massachusetts Institute of Technology (MIT) [Ross, 2006] en DYA [Wagter, 2005] waaruit het EAM raamwerk is ontwikkeld.

EAM wordt typisch beschreven als een opeenvolgend proces in vier stappen: (1) documenteer de huidige toestand, (2) beschrijf de gewenste toekomstige toestand, (3) ontwikkel het transitieplan en (4) implementeer het plan [Kotusev, 2017]. EA-artefacten op bedrijfsniveau beschrijven grotere reikwijdten in minder detail en zijn belangrijk voor architecten en managers op bedrijfsniveau. EA-artefacten op bedrijfsniveau kunnen bedrijfsdoelstellingen, principes, bedrijfsstructuren, kerndiagrammen, capaciteitskaarten en bedrijfsprocesmodellen en vele andere soorten artefacten zijn.

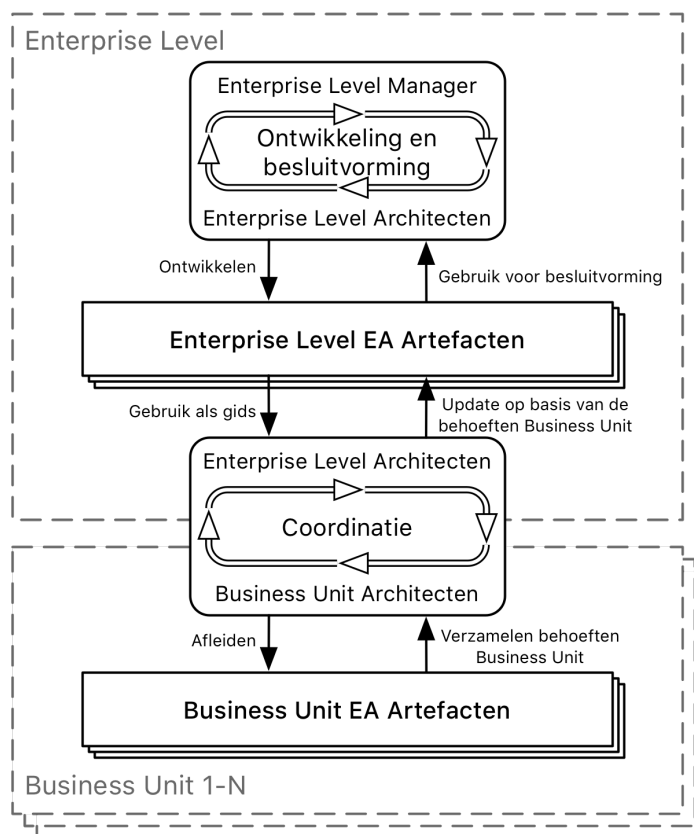
Business unit EA-artefacten beschrijven kleinere scopes in meer detail en zijn belangrijk voor business unit architecten en managers, maar ook voor IT-personeel. Net als bij EA-artefacten op bedrijfsniveau kunnen het verschillende soorten artefacten zijn, kunnen ze op elke redelijke manier systematisch worden opgebouwd en elke combinatie van huidige staten, toekomstige staten en routekaarten beschrijven. De reikwijdte van hun beschrijving kan variëren van onafhankelijke bedrijfstakken.

Het succesvol EAM houdt niet alleen in het maken van de noodzakelijke EA-artefacten, maar ook in het opzetten van de benodigde bedrijfsprocessen, zoals ontwikkelings-, besluitvorming-, coördinatie- en implementatieprocessen. Deze dragen allemaal bij aan het succes van de EA doelen [Kotusev, 2017]. In figuur '2.1-Vereenvoudigde conceptueel model van EAM voor gedecentraliseerde ondernemingen' is een deel van het volledige conceptuele model van EAM weergegeven, dat de interactie weergeeft tussen de enterprise level architectuur en de business unit architectuur. De directie Digitale Overheid (DO) bevindt zich op het enterprise level management en Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) verzorgt voor de coördinatie tussen de enterprise level architecten en de business unit architecten. Het EAM raamwerk is daarmee zeer geschikt voor dit onderzoek, omdat dit raamwerk is ontwikkeld uit de traditionele EA-raamwerken. Het volledige conceptueel model van EAM zoals Kotusev [Kotusev, 2017] heeft beschreven, is weergegeven in bijlage 'A-Verdieping', paragraaf A.2.1.

Het opzetten van de benodigde bedrijfsprocessen, zoals de besluitvormingsprocessen en de coördinatieprocessen, is een belangrijk onderwerpen binnen dit onderzoek. Deze processen kunnen niet los worden gezien van het gehele model, dus nemen we met name het deel dat is weergegeven in figuur 2.1 mee in het empirisch onderzoek.

Enterprise Architectuur BOLD

In het artikel van Lnenicka en Komarkova [Lnenicka, 2019] wordt het voorstel van de Government Enterprise Architecture Framework (GEAF) beschreven als geschikt voor het opslaan, verwerken en publiceren van BOLD met behulp van cloud computing. Het GEAF volgt twee benaderingen: multidimensionaliteit gebaseerd op de indeling in lagen voor cloud computing en de BOLD-levenscyclusfasen. De volgende belangrijke architectuurvisies worden voorgesteld: a) proces; b) functie; c) gegevens; d) software; en e) hardware. Beveiliging en



Figuur 2.1: Vereenvoudigde conceptueel model van EAM voor gedecentraliseerde ondernemingen [Kotusev, 2017, p. 16]

privacy omvatten alle standpunten en worden ook besproken met aandacht voor de behoeften van de publieke sector.

Het GEAF raamwerk ondersteunt het gebruik van BD en BOLD binnen de EA en is afgeleid van TOGAF [Jolley, 2016] en de Multidimensional Management and Development of Information Systems (MMDIS) [Bruckner, 2012] methodologie.

In dit onderzoek richten ons niet op de gegevens-, software- en hardware-architectuurvisies. Maar richten we ons meer op de procesarchitectuur, omdat daarmee een set regels, principes en richtlijnen kunnen worden vastgesteld. Daarmee nemen we de procesarchitectuur GEAF mee in het empirisch onderzoek.

2.3.3. Data Governance, Big Data en Open data

Over het algemeen verwijst governance naar de regels en praktijken waarmee de raad van bestuur ervoor zorgt dat strategieën worden ingevoerd, gecontroleerd en gerealiseerd [Rau, 2004]. Governance begint in eerste instantie op bedrijfsniveau, waar het een kader biedt om managers te ondersteunen bij hun dagelijkse activiteiten [Rau, 2004].

De opkomst van BD kan veel nieuwe kansen geven voor organisaties om waarde te creëren uit data. Om deze uitdagingen het hoofd te kunnen bieden, kunnen organisaties DG invoeren en gebruikmaken van een DG raamwerk [Baijens, 2020]. In het onderzoek van [Baijens, 2020] worden de vier leidende principes: *accountability*, *accessibility*, *community*, en *uniformity*, gebruikt als basis voor een DG raamwerk [Avery, 2015].

Het onderzoek van Rau [Rau, 2004] is interessant, omdat deze de ontwerpdoelstellingen, rollen en relaties beschrijven van DG.

In het onderzoek van Welle Donker en van Loenen [Welle Donker, 2017] worden zes element beschreven van een volwassenheidsraamwerk voor DG van publieke overheidsdata. Deze twee wetenschappers [Rau, 2004; Welle Donker, 2017] zijn ook door de beleidsmedewerkers van DO aangedragen voor dit onderzoek, zoals aangegeven in paragraaf '1.5-Motivatie / relevantie'.

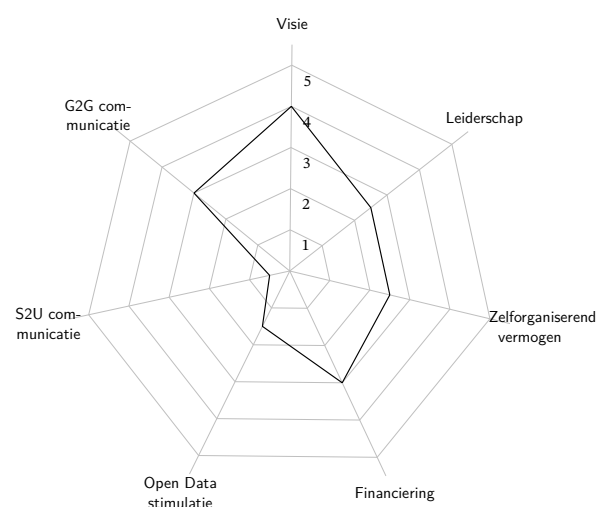
In het onderzoek van Khatri en Brown [Khatri, 2010] worden vijf beslissingsdomeinen van DG beschreven die belangrijk zijn voor een dataverwerkende organisatie. Hierin wordt de relatie tussen de IT-assets en data assets vergeleken, wat interessant kan zijn voor dit onderzoek [Khatri, 2010].

Het onderzoek van Al-Badi e.a. [AlBadi, 2018] beschrijft een BD Governance raamwerk, dat ondersteuning kan bieden bij beslissingen rondom BD door het toepassen van domeinspecifieke principes.

Volwassenheid governance open datavoorziening

Open data is momenteel een hot topic en hangt samen met het realiseren van ambities als een transparantere en efficiëntere overheid, het oplossen van maatschappelijke problemen en het vergroten van economische waarde [Welle Donker, 2017]. Ondanks hoge scores in deze beoordelingskaders, voldoet het daadwerkelijke (her)gebruik van Open Government Data (OGD) niet aan de verwachtingen. In figuur '2.2-Volwassenheidsniveau van open data governance in Nederland in 2014' zijn deze volwassenheid elementen weergegeven zoals deze zijn bepaald in Nederland in 2014.

De vijf elementen van de volwassenheid aangepast om de governance van open datavoorziening te bepalen volgens Welle Donker en van Loenen [Welle Donker, 2017] zijn:



Figuur 2.2: Volwassenheidsniveau van open data governance in Nederland in 2014 [Welle Donker, 2017]

1. Visie: een gemeenschappelijk doel bieden, een gefragmenteerde aanpak vermijden en de samenwerking tussen belanghebbenden stimuleren.
2. Leiderschap en controle: open data hebben een probleemeigenaar nodig die open data activiteiten stimuleert en coördineert. Bewustwording en capaciteitsopbouw kunnen leiden tot politiek draagvlak voor open data, wat een belangrijke succesfactor is, net als de werkvloer ondersteuning.
3. Communicatiekanalen: met wie, hoe en wat wordt er gecommuniceerd. In de beginfase zal dit vooral interne communicatie zijn en in latere fasen ook externe communicatie.
4. Zelforganiserend vermogen: de manier waarop aanbod aansluit op de vraag. In de beginfase zijn het veelal dataproviders die antwoorden op specifieke vragen nodig hebben en proactief open data promoten. In latere stadia zal het matchen van vraag en aanbod steeds meer een onderdeel worden van de organisatiecultuur.
5. Duurzame financiering: moet verder gaan dan de beginfase en ingebed worden in het budget van de organisatie voor gegevensbeheer en infrastructuur.

Dit onderzoek zal zich vooral richten op: (1) visie, (2) Leiderschap en controle, omdat deze een relatie hebben met principes en richtlijnen. De (3) communicatiekanalen hebben te maken met de domeinspecifieke principes en richtlijnen. Dat wil niet zeggen dat de andere topics van minder belang zijn, maar vallen niet binnen de scope van dit onderzoek.

Data Governance en Master Data Management

Binnen de literatuur worden de twee termen Master Data (MD) Management en Metadata Management naast elkaar gebruikt en hebben overlappende definities. MD bestaat uit vele datatypering, waaronder metadata, daarom gebruiken we binnen dit onderzoek de term Master Data Management (MDM).

In het onderzoek van Khatri en Brown [Khatri, 2010] worden vijf beslissingsdomeinen van DG beschreven: (1) *Data Principles*, (2) *Data Quality*, (3) *Metadata*, (4) *Data Access* en (5) *Data Lifecycle*. Waarbij de Data Principles bij de rol van Enterprise Data commissie hoort en de Metadata bij de rol van Enterprise data commissie en Enterprise data architecten. De taakstelling die gesteld wordt over het DG domein Data Principles is: "Clarifying the role of data as an asset" ([Khatri, 2010]). De taakstelling die gesteld wordt over het DG domein Metadata is: "Establishing the semantics or 'content' of data so that it is interpretable by the users" ([Khatri, 2010]).

Omdat metadata onderdeel is van MD en daarmee onderdeel van MDM, nemen we metadata van Khatri en Brown [Khatri, 2010] ook mee in dit onderzoek. Daarmee zijn zowel de taakstellingen van de (1) *Data Principles* als de richtlijnen voor het gebruik van (3) *Metadata* belangrijke onderwerpen om mee te nemen in het empirisch onderzoek.

2.3.4. FAIR Data Management, Big Data en Open data

In het onderzoek van Wilkinson e.a. [Wilkinson, 2016] worden de FAIR Data Principles beschreven voor wetenschappelijke (meta)data. Wat voor dit onderzoek van belang is of deze FAIR Data Principles ook kunnen worden toegepast op de publieke overheidsdata om deze beter vindbaar te maken en daarmee efficiënt te kunnen herbruiken.

In het onderzoek van Hasnain en Rebholz-Schuhmann [Hasnain, 2018] worden de FAIR Data Principles vergeleken met de vijf volwassenheidsniveaus van open data uit de figuur 'A.1-5 sterren stappen'.

In het onderzoek van Hasnain en Rebholz-Schuhmann [Hasnain, 2018] wordt geconcludeerd dat LOD van Berners-Lee [BernersLee, 2006] een duidelijk mandaat geeft aan de openheid van data. De FAIR Data Principles van [Wilkinson, 2016] stellen voor het ontwerpen van herbruikbaarheid dat er gebruik wordt gemaakt van duidelijke en toegankelijke licenties. Daarnaast verwijst FAIR Data Principles sterk naar de met een veelheid aan nauwkeurige en belangrijke attributen beschreven data en metadata die nodig is om hergebruik te verbeteren. Volgens de LOD-principes zouden dergelijke metagegevens ook als interoperabele gegevens worden beschouwd. De vereiste om gegevens uit te breiden met metagegevens, geeft wel aan dat FAIR Data Principles een uitbreiding is van de LOD. Hasnain en Rebholz-Schuhmann [Hasnain, 2018]

De leidende principes binnen FAIR volgens Wilkinson e.a. [Wilkinson, 2016, p. 4] zijn weergegeven in tabel 'A.5-FADAM FAIR raamwerk' in bijlage 'A-Verdieping'.

In het onderzoek van de Figueiredo e.a. [dFigueiredo, 2019] wordt een duidelijk verband gelegd tussen zowel de principes van het Data Management Framework (DAMA-DMBOK2) als de FAIR Data Principles. Hierbij wordt aangegeven dat zes van de acht DAMA-DMBOK-principes geen verband houden met FAIR, maar principes zijn die vanuit een algemeen organisatieperspectief worden bekeken waarbij het senior management betrokken is [dFigueiredo, 2019, ch. 4]. Dit heeft uiteindelijk geleid tot het voorstel tot het FADAM raamwerk van de Figueiredo e.a. [dFigueiredo, 2019], zie tabel 'A.5-FADAM FAIR raamwerk'.

Op basis van het onderzoek van de Figueiredo e.a. [dFigueiredo, 2019], Hasnain en Rebolz-Schuhmann [Hasnain, 2018] en Wilkinson e.a. [Wilkinson, 2016] kan worden geconcludeerd dat de FAIR Data Principles goed toepasbaar kunnen zijn om het hergebruik van de publieke overheidsdata te verbeteren. Daarnaast voegt het FADAM raamwerk de managementplanning en management commitment [dFigueiredo, 2019] aan de FAIR Data Principles. Hiermee is het FADAM raamwerk in combinatie met de FAIR Data Principles één van de belangrijkste onderwerpen voor het empirisch onderzoek.

2.3.5. Community Models, Big Data en Open data

In het onderzoek van Lyon, Alexander Ball e.a. [Lyon, 2012b] wordt het Community Capability Model Framework (CCMF) beschreven, dit raamwerk is mogelijk interessant binnen dit onderzoek bij beslissingen rond CM, door het toepassen van domeinspecifieke richtlijnen. In het onderzoek van Curry e.a. [Curry, 2010] worden enkele sociale en technische *Best Practices* aangegeven hoe het beste kan worden omgegaan met datacommunities. Misschien kunnen deze helpen bij het beantwoorden van de vragen van dit onderzoek.

Het CCMF van Lyon, Alexander Ball e.a. [Lyon, 2012b] is een hulpmiddel voor het evalueren van de huidige bereidheid van een gemeenschap om data-intensief onderzoek uit te voeren, en voor het identificeren van gebieden waar veranderingen moeten worden aangebracht om de capaciteit te vergroten. De CCMF is ontwikkeld met de betrokkenheid van een breed scala aan belanghebbenden en geïnteresseerde partijen. De onmiddellijke volgende stap zal zijn om het te valideren door het raamwerk toe te passen op enkele onderzoeksgemeenschappen [Lyon, 2012a; Lyon, 2012b].

In het CCMF worden acht capaciteitsfactoren beschreven, deze zijn (1) *collaboration*, (2) *skills and training*, (3) *openness*, (4) *technical infrastructure*, (5) *common practices*, (6) *economic and business models*, (7) *legal and ethical issues* en (8) *academic culture* [Lyon, 2012b].

Deze capaciteitsfactoren van CCMF kunnen de *Best Practices* van CM mogelijk ondersteunen.

De opkomst van pre-concurrentiële datasamenwerkingen is een belangrijke ontwikkeling: pre-concurrentiële data is informatie die kan worden gedeeld zonder dat dit een commercieel voordeel oplevert. Een voorbeeld hiervan is het Community Driven Data Curation (CDDC), waarbij de community meewerkt aan het verzamelen van data voor de organisatie. Een effectieve community-based data onderhoudsstrategie is sterk afhankelijk van de community die het bedient. Vroegtijdige betrokkenheid van de belangrijkste betrokkene, een continu communicatiekanaal voor de gemeenschap, duidelijke prikkels en een effectief bestuursmodel zijn belangrijke sociale aspecten voor gemeenschapontwikkeling [Curry, 2010].

In het CCMF zijn geen aantoonbare relaties gevonden met de principes en richtlijnen van EA en of DG. Daarnaast zijn er ook niet veel wetenschappelijke artikelen over dit onderwerp te vinden, zie tabel '2.1-Uitvoering literatuuronderzoek'.

Het CCMF en CDDC kunnen mogelijk richting geven aan domeinspecifieke richtlijnen. Bijvoorbeeld door de acht beschreven capaciteitsfactoren van CCMF verder te ontwikkelen. Daarmee nemen we het CCMF mee in het empirisch onderzoek.

2.3.6. Beantwoording literatuurvragen

De vier overgebleven vragen L1 tot en met L5 die in deze literatuurstudie beantwoord kunnen worden zijn schematisch samengevat in tabel 'A.8-Beantwoording literatuurvragen' in bijlage 'A-Verdieping'. In deze tabel is

het gevonden wetenschappelijk onderzoeksveld weergegeven met de gevonden principes en richtlijnen. Als eerste wordt aangegeven welke literatuurvraag met de gevonden principes en richtlijnen beantwoord kan worden. Alleen de eerste keer wordt de volledige literatuurvraag weergegeven, daarna alleen een verwijzing. De principes en richtlijnen zijn gelabeld en genummerd.

De labels van tabel A.8 bestaan uit:

- Een wetenschappelijk vakgebied die bestaan uit één of meerdere vakgebieden waarop de te beschrijven principe of richtlijn betrekking heeft.
- Een uniek volgnummer binnen de combinatie van vakgebieden.

Hiermee zijn ook de vijf resterende literatuur vragen ook beantwoord.

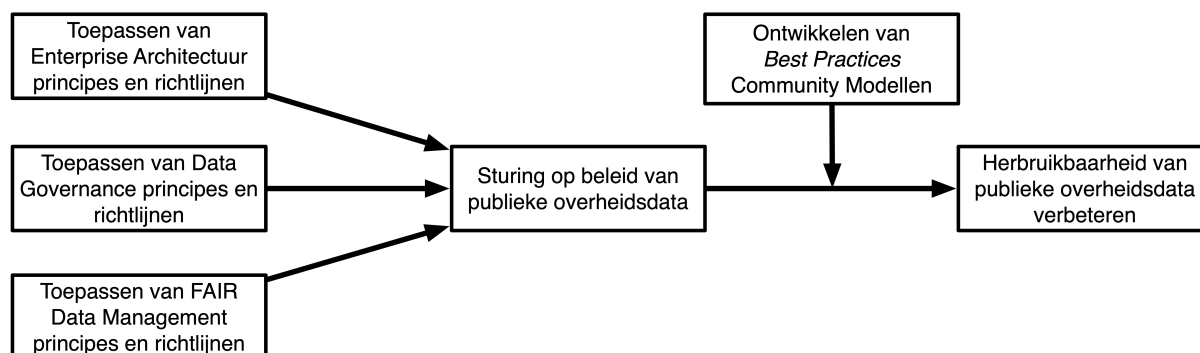
2.3.7. Conceptueel model

Het Conceptueel model van dit onderzoek bestaat uit vijf kernbegrippen die invloed op elkaar hebben zoals in figuur '2.3-Conceptueel model' is weergegeven [Verschuren, 2015, pp. 265–301]. Zo hebben de principes en richtlijnen van EA, DG en FADAM en het toepassen hiervan een directe invloed op de sturing van het beleid van publieke overheidsdata. De sturing op het beleid van publieke overheidsdata heeft een directe invloed op de herbruikbaarheid van de publieke overheidsdata. Dus door het toepassen van de principes en richtlijnen van EA, DG en FADAM kan indirect de herbruikbaarheid van de publieke overheidsdata worden verbeterd door sturing op het beleid van publieke overheidsdata.

Daarnaast hebben ook de *Best Practices* van CM een invloed op de herbruikbaarheid van de publieke overheidsdata. De herbruikbaarheid van de publieke overheidsdata wordt niet direct beïnvloed, maar heeft invloed op bepaalde gebruikersgroepen van publieke overheidsdata. Deze invloed kan worden versterkt door meer gebruik te maken van *Best Practices* van CM. In tabel '2.2-Conceptueel model' is aangegeven welke dimensies antwoord geven op de onderzoeksvragen voor het empirisch onderzoek, in tabel 'A.9-Kernbegrippen, dimensies en indicatoren' zijn de verschillende dimensies en kernbegrippen in detail uitgewerkt.

Beantwoording onderzoeksvragen voor het empirisch onderzoek	
E1.	Welke algemene principes en richtlijnen van EA en DG worden er gebruikt om te helpen om de sturing op het beleid rond het hergebruik van publieke overheidsdata te verbeteren?
E2.	Welke principes van FADAM worden er gebruikt om te helpen om de sturing op het beleid rond het hergebruik van publieke overheidsdata te verbeteren?
E3.	Welke <i>Best Practices</i> van CM worden er gebruikt om te helpen om het hergebruik van publieke overheidsdata te verbeteren?

Tabel 2.2: Beantwoording onderzoeksvragen voor het empirisch onderzoek



Figuur 2.3: Conceptueel model

2.4. Doel van het vervolgonderzoek

De literatuurstudie heeft een beter begrip en inzicht gegeven door belangrijk wetenschappelijk onderzoek op het gebied van EA, DG, FADAM en CM te bestuderen en te analyseren. Hierbij is gekeken naar algemene principes en de vertaling hiervan naar beleidslijnen [Josey, 2016; OU, 2009; vdBerg, 2004]. Daarnaast is gekeken naar domeinspecifieke architectuurprincipes en richtlijnen op de bovengenoemde begrippen [Henderson, 2017; Josey, 2016; OU, 2009]. Hierdoor zijn enkele richtlijnen en principes gevonden die de sturing op het beleid van publieke overheidsdata om het hergebruik te verbeteren. Daarnaast zijn er enkele *Best Practices* gevonden die de herbruikbaarheid van publieke overheidsdata verder kunnen verbeteren.

Het is belangrijk om te weten hoe deze richtlijnen en principes in de praktijk gemeten kunnen worden, die ook wel dimensies worden genoemd [Doorewaard, 2019, p. 55]. De drie meest voorkomende bronnen van informatie in het praktijkgericht kwalitatieve onderzoek zijn volgens Hans Doorewaard e.a. [Doorewaard, 2019, p. 94]: interviews, documenten en observatie. In Saunders e.a. [Saunders, 2019, Ch. 8, Ch. 10] worden interviews, documenten genoemd als mogelijke informatiebronnen, en deze zullen ook worden gebruikt voor het empirisch onderzoek. In tabel 2.3 is bepaald welke van de informatiebronnen mogelijk zijn voor de gegevensverzameling van de gevonden dimensies. Hierbij is alleen gekeken naar de theorie en welke methode van informatieverzameling hierbij de beste kwalitatieve informatie kan opleveren. In hoofdstuk 3 zal worden gekeken wat voor dit onderzoek de mogelijkheden zijn.

Informatie bronnen	Interviews	Documentenanalyse
Principes, richtlijnen en Best Practices		
EA-1 en EA-2	Beste methode, omdat het vaak om een relatief kleine groep functionarissen gaat met diepgaande kennis en ervaring.	Kan goed worden gebruikt voor triangulatie.
EA-3 en EA-4	Is mogelijk, maar tijdrovend vanwege de vaak relatief grote groep functionarissen met verspreide kennis en ervaring.	Mogelijk als triangulatie te gebruiken als bekend is waar deze informatie gevonden kan worden.
EA-5 en EA-6	Kan mogelijk via interviews worden achterhaald, maar kan worden gebruikt als triangulatie.	Dit is lastig uit documenten te achterhalen.
EA-BD-1 T/M EA-BD-6	Is mogelijk, maar tijdrovend vanwege de vaak relatief grote groep functionarissen met verspreide kennis en ervaring.	Mogelijk als triangulatie te gebruiken als bekend is waar deze informatie gevonden kan worden.
DG-BD-1 T/M DG-BD-5	Beste methode, omdat het vaak om een relatief kleine groep functionarissen gaat met diepgaande kennis en ervaring.	Dit is lastig uit documenten te achterhalen.
DG-BD-6 DG-BD-7	Is mogelijk, maar tijdrovend vanwege de vaak relatief grote groep functionarissen met verspreide kennis en ervaring.	Mogelijk als triangulatie te gebruiken als bekend is waar deze informatie gevonden kan worden.
FADAM-BD-1 en FADAM-BD-2	Beste methode, omdat het vaak om een relatief kleine groep functionarissen gaat met diepgaande kennis en ervaring.	Dit is lastig uit documenten te achterhalen.
FADAM-BD-3 T/M FADAM-BD-6	Is mogelijk, maar tijdrovend vanwege de vaak relatief grote groep functionarissen met verspreide kennis en ervaring.	Dit is lastig uit documenten te achterhalen.
CM-BD-1 T/M CM-BD-12	Lastig en tijdrovend vanwege de vaak relatief grote groep functionarissen met verspreide kennis en ervaring.	Lastig en tijdrovend dit middels documenten te achterhalen.

Tabel 2.3: Kwalitatieve informatiebronnen voor de verschillende dimensies

2.4.1. Vervolgonderzoek

In het vervolgonderzoek wordt gekeken naar de mogelijkheid om de gevonden dimensies te kunnen meten in de praktijk. Hierbij wordt het globaal conceptueel model verder uitgewerkt tot het conceptueel ontwerp en het onderzoeksmodel van figuur '3.1-Onderzoeksmodel'. In het empirisch onderzoek wordt gekeken of deze inzichten in de werkelijkheid ondersteund worden [Doorewaard, 2019].

3

Methodologie

Als onderzoeksmethode is gekozen voor een deductieve kwalitatieve benadering door gebruik te maken van een conceptueel model [Doorewaard, 2019, pp. 14-15]. Eerst wordt conceptueel ontwerp gemaakt, paragraaf 3.1, waarna deze wordt uitgewerkt tot een technisch ontwerp, paragraaf 3.2. Daarna wordt verder ingegaan op de gegevensanalyse, paragraaf 3.3 en als laatste de reflectie, paragraaf 3.4.

3.1. Conceptueel ontwerp: keuze van onderzoeksmethode(n)

Als onderzoeksstrategie voor het vervolgonderzoek is gekozen voor een enkelvoudige casusstudie [Doorewaard, 2019; Saunders, 2019]. De casus-organisatie is de Nederlandse rijksoverheid, specifiek binnen het Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK) die verantwoordelijk is voor het beleid rond publieke overheidsdata. Als onderzoeksparadigma is gekozen voor een interpretatief paradigma op basis van Saunders e.a. [Saunders, 2019, pp. 138-141]. Dit houdt in dat dit onderzoek rekening houdt met een subjectieve onderzoeksfilosofie binnen een reguleringsperspectief.

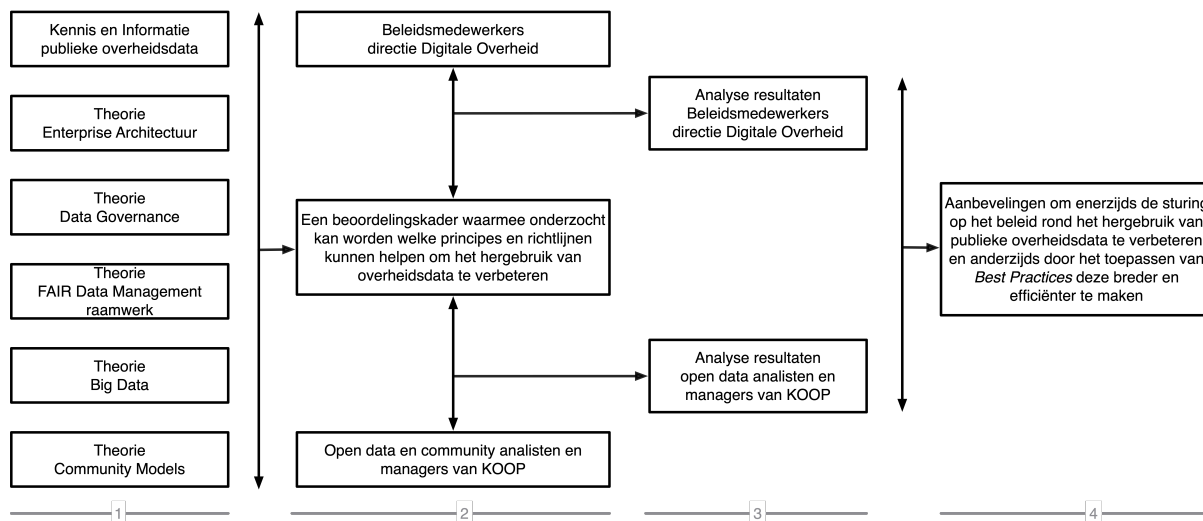
Voor een kwalitatief onderzoeksontwerp zijn er enkele onderzoekstechnieken bruikbaar, zoals observatie en interviews [Saunders, 2019, pp. 179-180]. In paragraaf 2.4 zijn de onderzoeksdimensies bepaald en hierbij is gekeken naar de mogelijke informatiebronnen. Voor bijna alle dimensies zijn interviews de meest voor de hand liggende informatiebron. Als tweede komt de documentenanalyse naar voren, maar deze is niet geschikt voor een kwalitatief onderzoeksontwerp volgens Saunders e.a. [Saunders, 2019]. Voor de meeste dimensies gelden dat observaties de minst geschikte informatiebron is, ook omdat deze veel tijd kosten. Daarom is voor het empirisch onderzoek gekozen om semigestructureerde interviews te doen.

De interviews zullen worden gehouden met enkele Beleidsmedewerkers van directie Digitale Overheid (DO), die verantwoordelijk zijn voor het ontwikkelen van beleid rond publieke overheidsdata. Daarnaast zullen enkele publieke overheidsdata architecten van Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) die verantwoordelijk zijn voor het omzetten van beleid in concrete richtlijnen voor publieke overheidsdata [KOOP, 2021b].

3.1.1. Onderzoeksmodel

Het onderzoeksmodel van dit onderzoek is weergegeven in figuur '3.1-Onderzoeksmodel'. In de eerste fase is de theoriestudie gedaan van de verschillende onderzoeksdomeinen, Enterprise Architectuur (EA), Data Governance (DG), FAIR Data Management (FADAM), Big Data (BD) en Community Models (CM) gecombineerd met algemene kennis en informatie van publieke overheidsdata. Hieruit zijn enkele principes en richtlijnen naar voren gekomen die mogelijk invloed kunnen hebben op de sturing op beleid van publieke overheidsdata. Daarnaast zijn er enkele *Best Practices* van CM naar voren gekomen die invloed hebben op de doelmatigheid van dit beleid. Door het toepassen van deze *Best Practices* kan de sturing op beleid van publieke overheidsdata het hergebruik van publieke overheidsdata nog breder en efficiënter maken.

In de tweede fase worden deze dimensies gecontroleerd door semigestructureerde interviews bij de twee hoofdgroepen bestaande de Beleidsmedewerker van DO en open data en communities analisten van KOOP. Vervolgens in de derde fase worden de resultaten van beide hoofdgroepen apart van elkaar geanalyseerd. In de vierde fase worden deze resultaten samengevoegd en kunnen er mogelijke aanbevelingen worden gedaan om de sturing op beleid rond het hergebruik van publieke overheidsdata te verbeteren.



Figuur 3.1: Onderzoeksmodel

3.2. Technisch ontwerp: uitwerking van de methode

Als onderzoeksmethode voor dit onderzoek is gekozen om interviews te houden met Beleidsmedewerkers van DO en analisten en managers van KOOP van BZK. Voor het empirisch deel van het onderzoek zal een vergelijkbare structuur worden aangehouden als is gebruikt in het vooronderzoek zoals in bijlage 'B-Vooronderzoek' is gebruikt. Daarmee is één casus-organisatie als geheel onderzocht, dit staat ook bekend als een *holistic casestudy* [Yin Robert, 2017]. Als onderzoeksmethode voor dit kwalitatieve onderzoek is gekozen om interviews te houden binnen de casus-organisatie binnen twee verschillende afdelingen [Doorewaard, 2019, p. 94].

Er is gekozen om semigestructureerde interviews te houden om dit probleemgebied te onderzoeken [Saunders, 2019, p.436-444]. De resultaten van deze interviews worden eventueel gecombineerd met aangedragen secundaire documentatie.

3.2.1. Interviews

Tijdens de semigestructureerde interviews zal de geïnterviewde gemotiveerde antwoorden geven op de vragen. Hierbij is er ruimte voor verdieping en verbreding, wat mogelijk kan leiden tot bijvangst. De geanonimiseerde getranscribeerde interviews zijn, waar mogelijk, beschikbaar voor vervolgonderzoek en kunnen worden gebruikt voor statistische analyses.

Als template voor het transcriberen van interviews wordt de template van paragraaf 'C.1.1-Samenvatting transcriptie interview' gebruikt. Dit template wordt in het empirisch deel van het onderzoek verder aangevuld met de gegevens uit tabel 'A.9-Kernbegrippen, dimensies en indicatoren', in overeenstemming met de kernbegrippen van het conceptueel model figuur 2.3.

Er wordt tijdens de interviews actief gevraagd naar documentatie waarmee triangulatie kan worden gedaan om daarmee de validiteit en betrouwbaarheid van dit onderzoek te vergroten.

3.2.2. Bronnen

In het onderzoek van Kotusev [Kotusev, 2017] zijn er verschillende managers beschreven binnen het Enterprise Architecture Management (EAM) raamwerk zoals weergegeven in figuur 2.1. Op het enterprise niveau, wordt de enterprise level manager beschreven die besluitvorming ontwikkelt voor de gehele organisatie. Voor het hergebruik van publieke overheidsdata wordt dit gedaan door de Beleidsmedewerkers van DO binnen BZK.

De coördinatie tussen enterprise level en business unit level door enerzijds enterprise level architecten en anderzijds business unit architecten. Binnen dit onderzoek is gekozen om alleen op enterprise niveau te kijken en niet op business unit niveau. Binnen BZK is KOOP verantwoordelijk voor het invullen van deze coördinatie. Het ontwikkelen van de enterprise level EA artefacten wordt gedaan door onder andere de , Community manager en Datamanager en analisten van KOOP.

3.2.3. Plan van aanpak

Voor het empirisch deel van het onderzoek is een eerste plan van aanpak gemaakt. Deze kan mogelijk in het empirisch deel nog worden aangepast, maar voor nu in schets opgenomen, zie tabel '3.1-Plan van aanpak'.

Stap	Acties
1. Voorbereiding	<ul style="list-style-type: none">• Contact leggen DO en KOOP• Afspraken vastleggen• Opstellen interview vragen
2. Eerste interview DO	<ul style="list-style-type: none">• interview afspraak maken• interview afnemen• antwoorden analyseren
3. Evaluatie eerste interview DO	<ul style="list-style-type: none">• evaluatie interview• evaluatie antwoorden• bijsturen interview vragen
4. Volgende interviews DO	<ul style="list-style-type: none">• interview afspraak maken• interview afnemen• antwoorden analyseren
5. Eerste interview KOOP	<ul style="list-style-type: none">• interview afspraak maken• interview afnemen• antwoorden analyseren
6. Evaluatie eerste interview KOOP	<ul style="list-style-type: none">• evaluatie interview• evaluatie antwoorden• bijsturen interview vragen
7. Volgende interviews KOOP	<ul style="list-style-type: none">• interview afspraak maken• interview afnemen• antwoorden analyseren
8. Analyseren van de gegevens	<ul style="list-style-type: none">• gegevens analyseren• samenvattingen valideren
9. Resultaten verwerken in verslag	<ul style="list-style-type: none">• conclusie trekken• aanbevelingen maken

Tabel 3.1: Plan van aanpak

3.3. Gegevensanalyse

De semigestructureerde interviews worden vastgelegd door geluidsopnames eventueel gecombineerd met beeldopnames. Deze kwalitatieve data worden omgezet middels kwalitatief transcriberen naar tekst en wordt daarna gevalideerd door de geïnterviewde. Vervolgens wordt hiervan een samenvatting gemaakt door thematische analyse toe te passen of door de gearde theorie methode toe te passen [Saunders, 2019, p.668-674]. Ook het eindresultaat hiervan wordt geverifieerd door de geïnterviewde. Deze methode is ook gebruikt in het vooronderzoek die te vinden is in bijlage B.

Het nadeel van transcriberen is de factor tijd, maar heeft als voordeel een gedwongen gedegen analyse van verkregen kwalitatieve data.

Een gerichte documentenanalyse zou kunnen worden toegevoegd om daarmee een extra triangulatie te kunnen doen van de informatie uit de interviews. Dat zal niet voor alle dimensies kunnen, maar tabel '2.3-Kwalitatieve informatiebronnen voor de verschillende dimensies' zou daar bij als basis dienen.

Als template voor het analyseren van de kwalitatieve data wordt de template van paragraaf 'C.2-Template analyse documentatie' gebruikt. Dit template wordt in het empirisch deel van het onderzoek verder ingevuld met de gegevens uit tabel 'A.9-Kernbegrippen, dimensies en indicatoren'. Door het analyseren van kwalitatieve data met de indicatoren van tabel A.9 kan het template van paragraaf D.2 worden ingevuld.

3.4. Reflectie t.a.v. validiteit, betrouwbaarheid en ethische aspecten

Inherent aan casestudy onderzoek wat uitgevoerd wordt om een fenomeen binnen de *real life* omgevingen te onderzoeken is de problematiek met onder andere de validiteit en betrouwbaarheid [Yin Robert, 2017]. Het uitvoeren van kwalitatief onderzoek brengt enkele methodologische uitdagingen met zich mee rondom betrouwbaarheid en generaliseerbaarheid van het onderzoek [Saunders, 2019]. In deze paragraaf wordt de kwaliteitseis validiteit beschreven op basis van de criteria: constructvaliditeit, interne validiteit, externe validiteit, betrouwbaarheid en ethische kwesties.

3.4.1. Constructvaliditeit

De constructvaliditeit volgens Yin Robert [Yin Robert, 2017] behelst de vraag of in de onderzoeksvragen gedefinieerde begrippen op de juiste en correcte wijze gemeten worden. Het onderzoek vindt plaats binnen een *holistic casestudy*, waardoor de triangulatie wordt bemoeilijkt. Door meerdere menselijke bronnen te gebruiken wordt toch enige vorm van triangulatie toegepast. Dit kan verder worden versterkt door het analyseren van beschikbaar gestelde documenten door de geïnterviewden.

3.4.2. Interne validiteit

De interne validiteit behelst het aantoonbaar maken dat de gevonden verklaringen correct zijn. Omdat er sprake is van verkennend onderzoek is het in dit onderzoek niet mogelijk om resultaten te toetsen op causale verbanden. Binnen dit onderzoek zal worden geprobeerd om vanuit de theorie verklaringen te geven om bias te voorkomen.

Om interview bias te voorkomen wordt aangegeven dat het interview vertrouwelijk is en wordt geanonimiseerd. Interviewer bias zal worden voorkomen door de techniek: vragen, luisteren, samenvatten en doorvragen, waarbij doorvragen in lijn ligt met de voorbereide interviewvragen.

3.4.3. Externe validiteit

Met externe validiteit wordt de generaliseerbaarheid van onderzoeksresultaten naar vergelijkbare cases. Omdat het onderzoek een *single holistic casestudy* is daarmee de generaliseerbaarheid beperkt. Het onderzoek is deels gestoeld op internationale wetenschappelijk onderzoek waardoor generalisatie toch beperkt mogelijk is.

3.4.4. Betrouwbaarheid

Dit onderzoek wordt uitgevoerd binnen de beschreven normen de bepalingen van "De Nederlandse Gedragscode Wetenschapsbeoefening". Deze stelt dat binnen wetenschappelijk onderzoek er voldaan moet worden aan de principes: eerlijkheid, zorgvuldigheid, transparantie, onafhankelijkheid en verantwoordelijkheid [Algra, 2018].

3.4.5. Ethische kwesties

Bij het uitvoeren van onderzoek spelen ethische kwesties ook een rol. Een onderzoeker dient moreel besef te hebben bij de uitvoering van zijn of haar onderzoek. Ethische kwesties kunnen te maken hebben met verschillende aspecten [Saunders, 2019]:

- Privacy van de deelnemers aan het onderzoek.
- Het vrijwillige karakter van deelname.
- Toestemming van deelnemers.
- Vertrouwelijkheid van de data en anonimiteit van de deelnemers.
- Reacties van deelnemers zoals schaamte en stress.
- Effect van deelnemers op de manier van rapporteren.
- Gedrag en objectiviteit van de onderzoeker.

Extra aandacht voor de ethische issues is nodig bij dit onderzoek vanwege de persoonlijke benadering bij een kwalitatief onderzoek. Daarnaast vereist nieuwe wetgeving over dataprotectie een gereserveerde houding bij het verwerken van gegevens. De privacy wordt voor dit onderzoek gewaarborgd door de namen van de geïnterviewden niet te gebruiken, maar alleen de functie die ze vervullen binnen de casus-organisatie. De casus-organisatie, Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties heeft een publieke functie en hecht waarde aan openheid, ook met betrekking tot dit onderzoek.

Tijdens het onderzoek worden aan alle methodologische uitdagingen aandacht besteed.

4

Resultaten

In dit hoofdstuk worden de resultaten van het onderzoek beschreven en de context van het empirische onderzoek wordt toegelicht. Als eerste wordt het proces van het empirische onderzoekproces beschreven. Daarna worden de resultaten van de gehouden interviews en de analyse van de documenten besproken. Dit wordt gedaan om de resultaten per variabele in het conceptueel model van paragraaf '2.3.7-Conceptueel model' te kunnen presenteren in overeenstemming met de onderzoeksvragen.

4.1. Beschrijving onderzoeksproces

Deze paragraaf gaat kort in op de casus-organisatie. Er wordt besproken welke onderzoeksmethode is gebruikt en hoe de uitvoering ervan is verlopen. Daarnaast worden de geselecteerde respondenten beschreven, evenals hoe er vanuit de theorie naar het onderzoeksontwerp is gekomen. Tot slot wordt de aanpak voor de analyse van de verzamelde data besproken.

4.1.1. Onderzoeksmethode

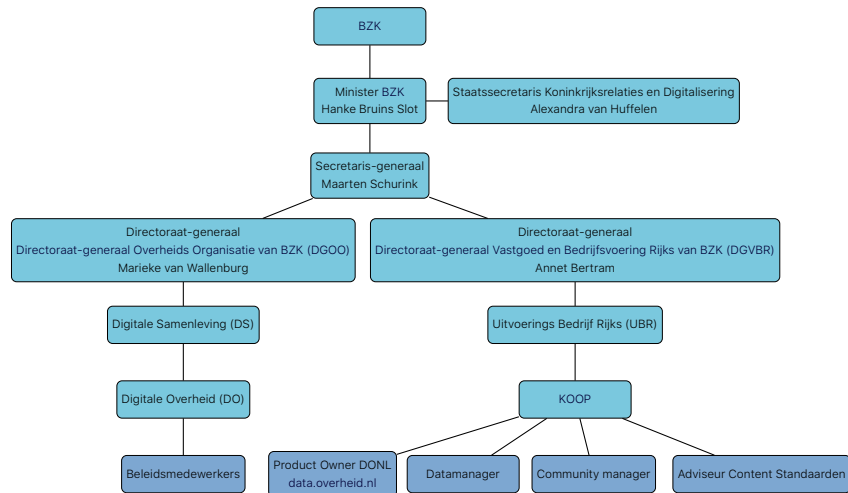
Het onderzoek is uitgevoerd zoals beschreven in paragraaf '3.2-Technisch ontwerp: uitwerking van de methode' en heeft plaatsgevonden op het beleidsniveau binnen de Nederlandse rijksoverheid. Met deze deductieve kwalitatieve methode is het mogelijk om inzicht te krijgen in hoe de sturing op het beleid van publieke overheidsdata vorm krijgt, waardoor we de onderzoeksvragen kunnen beantwoorden zoals aangegeven in tabel '2.2-Conceptueel model'. Er is op een paar punten afgeweken van het technisch ontwerp zoals beschreven in paragraaf 3.2, ten eerst zijn een aantal stappen in het plan van aanpak anders uitgevoerd. Hierbij is alleen het eerste interview geëvalueerd en hieruit bleek dat de interviewtijd beperkt was, maar dat er inhoudelijk niet veel behoefde te veranderen. Door strak aan de interviewvragen te houden en randzaken te beperken tot alleen die voor dit onderzoek relevant waren, was het mogelijk alle vragen beantwoord te krijgen. Ten tweede kwam er naar voren tijdens het eerste interview dat er mogelijk nog een mogelijk een interessante functionaris was voor dit onderzoek. Deze functionaris was niet in scope vanuit het vooronderzoek, maar geeft invulling aan het beleid rond open data standaarden en daarmee interessant voor dit onderzoek. Dit was de Adviseur Content Standaarden van Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) en deze is toegevoegd aan de lijst met te interviewen respondenten.

4.1.2. Introductie casus-organisatie

De Nederlandse rijksoverheid staat voor grote uitdagingen zoals de actieve openbaarmaking van overheidsinformatie. Deze actieve openbaarmaking is van groot belang voor de controleerbaarheid van de overheid en de herbruikbaarheid van overheidsinformatie door burgers en bedrijven [Koninkrijksrelaties, 2019]. Als casus-organisatie is gekozen voor het Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK), omdat dit verantwoordelijk is voor de wettelijke en beleidsmatige kaders van de digitale overheid. Binnen BZK

is directie Digitale Overheid (DO) verantwoordelijk voor de beleidsmatige kaders, en het KOOP ontwikkelt de standaarden en zorgt voor afstemming tussen de verschillende organisaties. Dit wordt onder andere gedaan in de Data Agenda Overheid (DAO). Daarnaast is KOOP ook verantwoordelijk voor het publiceren van publieke overheidsdata en alle standaarden die hiervoor van toepassing zijn. KOOP is de drijvende kracht achter www.overheid.nl en daarmee ook voor data.overheid.nl waar alle publieke overheidsdata te vinden is [KOOP, 2021b; KOOP, 2021c].

De afdeling DO valt onder het DGOO en bestaat uit verschillende Beleidsmedewerkers, elk met hun eigen specialisme. De afdeling KOOP valt onder DGVBR en bestaat uit veel verschillende disciplines, waaronder verschillende Datamanagers en Product Owner DONLs. In het organigram van figuur '4.1-Organigram casus-organisatie' is te zien hoe de verschillende functionarissen zijn geplaatst binnen BZK, deze zijn met een andere kleur aangegeven.



Figuur 4.1: Organigram casus-organisatie

Voor dit onderzoek zijn de Beleidsmedewerkers van DO geselecteerd die actief betrokken zijn bij het maken van beleid op het gebied van publieke overheidsdata. Ook zijn de Datamanager en Product Owner DONL van KOOP geselecteerd die betrokken zijn bij het ontwikkelen van standaarden op het gebied van publieke overheidsdata. Tijdens de interviewfase is de Adviseur Content Standaarden van KOOP toegevoegd, omdat dit de kwaliteit van het onderzoek zou kunnen verbeteren.

4.1.3. Uitvoering onderzoek

Uit het theoretisch kader van hoofdstuk '2-Theoretisch kader' zijn een aantal kernwaarden van Enterprise Architectuur (EA), Data Governance (DG), FAIR Data Management (FADAM) en Community Models (CM) in relatie tot Big Data (BD) naar voren gekomen die de herbruikbaarheid van publieke overheidsdata kunnen verbeteren. Na het beantwoorden van de literatuurvragen zijn er een aantal kernwaarden naar voren gekomen die gebruikt zijn om het conceptueel model van figuur '2.3-Conceptueel model' op te stellen. Hierna zijn de onderzoeksvragen vastgesteld voor het empirische onderzoek, deze zijn uitgewerkt in tabel 'C.1-Template transcriptie interview'. Op basis van onafhankelijke en modererende variabelen in conceptueel model is er een verdeling gemaakt van welke functionaris bepaalde interviewvragen zou kunnen beantwoorden. Hierdoor is er een beperkte vorm van constructvaliditeit gerealiseerd, om daarmee de betrouwbaarheid van de kwalitatieve data te vergroten. Om de data van de interviews te valideren, wordt er naar aanvullende documentatie gevraagd. Elk document wordt als secundaire kwalitatieve data vastgelegd in het template van tabel 'C.2-Template analyse document'.

De kwalitatieve data van de interviews worden vervolgens omgezet naar tekst door middel van kwalitatief transcriberen [Saunders, 2019, p.644-647]. Hiervoor is er gebruik gemaakt van ATLAS.ti, om audio en de getranscribeerde tekst te kunnen koppelen. Tegelijkertijd worden deze teksten voorzien van codering per dimensie en indicator. Per geïnterviewde worden deze teksten geëxporteerd om te kunnen worden gevalideerd door de geïnterviewden.

Daarna zijn de teksten per indicator overgenomen in het template van tabel 'C.2-Template analyse document'. Deze kwalitatieve tekst wordt vervolgens verder geanalyseerd door de across-cases analyse [Doorewaard,

2019, p.86, 135–139] toe te passen om per dimensie een samenvatting te verkrijgen. Deze kwalitatieve samenvatting is te vinden in bijlage 'D-Kwalitatieve interview data', paragraaf D.1.

De secundaire data zijn geanalyseerd volgens de methode van Saunders e.a. [Saunders, 2019, p.338-368], waarbij de algehele geschiktheid van data wordt bepaald. Deze secundaire data zijn te vinden in bijlage 'D-Kwalitatieve interview data', paragraaf D.2. De geschikte secundaire data worden vervolgens nauwkeuriger geanalyseerd en de betrouwbaarheid en validiteit wordt bepaald.

Alle verzamelde gegevens uit paragraaf D.1 en paragraaf D.2 worden per dimensie samengevat volgens de methoden van [Doorewaard, 2019, p.143-153]. De resultaten hiervan worden weergegeven in paragraaf '4.2.2-Resultaten documentanalyse'. Er zijn geen afwijkingen opgetreden ten opzichte van het oorspronkelijke plan uit hoofdstuk '3-Methodologie'.

4.1.4. Selectiecriteria respondenten

In dit onderzoek worden er interviews afgenomen met twee coördinerende Beleidsmedewerkers van de afdeling DO en Community manager, Datamanager en de Product Owner DONL van KOOP. Hiervoor zijn twee Beleidsmedewerkers van DO geselecteerd om de interviewvragen te beantwoorden met betrekking tot de sturing op het beleid van publieke overheidsdata en het ontwikkelen van *Best Practices* voor CM. Daarnaast zijn een Datamanager, een Adviseur Content Standaarden, een Community manager en een Product Owner (PO) van KOOP geselecteerd om de interviewvragen te beantwoorden met betrekking over het maken van het beleid rondom publieke overheidsdata en het implementeren van de *Best Practices* voor CM.

De verdeling van interviewvragen over de verschillende respondenten is weergegeven in tabel '4.1-Respondenten en onderzoeksvragen'. Dit is gedaan omdat elke respondent een specifieke functie heeft en bekend is met de bijbehorende kennis en ervaring, en daardoor niet bekend is met alle dimensies van dit onderzoek. Bovendien heeft de beschikbare tijd tijdens het interview invloed gehad op het aantal vragen dat gesteld kon worden, waardoor het aantal beperkt moest zijn.

Kernbegrip	Dimensie	Beleidsmedewerker een	Beleidsmedewerker twee	Product Owner DONL	Datamanager	Community manager	Adviseur Content Standaarden
Toepassen van Enterprise Architectuur principes en richtlijnen	EA	-	EA-1, EA-3	EA-1 t/m 6	EA-1 t/m 6	-	EA-1 t/m 6
	EA-BD	-	-	EA-BD-1 t/m 6	EA-BD-1 t/m 6	-	EA-BD algemeen
Toepassen van Data Governance principes en richtlijnen	DG-BD	DG-BD-1 t/m 6	DG-BD-1 t/m 6	-	DG-BD-6, DG-BD-7	-	DG-BD-1 t/m 6
Toepassen van FAIR Data Management principes en richtlijnen	FADAM-BD	FADAM-BD-1 t/m 2	FADAM-BD-1 t/m 2	FADAM-BD-1 t/m 6	FADAM-BD-1 t/m 6	-	FADAM-BD-1 t/m 6
Ontwikkelen van Best Practices Community Models	CM-BD	-	-	CM-BD-1 t/m 12	-	CM-BD-1 t/m 12	CM-BD-1 t/m 12

Tabel 4.1: Respondenten en onderzoeksvragen

4.1.5. Van theorie naar onderzoeksopzet

In dit onderzoek is het conceptueel model van figuur '3.1-Onderzoeksmodel' vastgesteld op basis van principes en richtlijnen die in de literatuur zijn gevonden voor EA, DG en FADAM, samen met de *Best Practices* voor CM. In hoofdstuk '3-Methodologie' is een technisch ontwerp gemaakt waarin de onderzoeksmethode is vastgelegd op basis van de kernbegrippen die zijn gevonden in de literatuur. Om de kernbegrippen in de praktijk te valideren, zijn de belangrijkste indicatoren per dimensie verzameld en zijn er per indicator interviewvragen geformuleerd om de kwalitatieve data te verkrijgen. Hiermee kunnen de indicatoren in de praktijk worden gevalideerd.

In de interviewtemplate van paragraaf 'C.1-Interview template' en het template voor het analyseren van aanvullende documentatie uit paragraaf 'C.2-Template analyse documentatie', wordt steeds teruggekeken naar het conceptueel model. Het doel is om met behulp van deze middelen antwoorden te krijgen op de deelvragen en uiteindelijk het conceptueel model te bevestigen. Om extra uitleg te kunnen geven aan de in de kernbegrippen die in de literatuur zijn gevonden, wordt gevraagd om op een Likert-schaal van vijf punten aan te geven hoe toepasbaarheid het kernbegrip is voor het verbeteren van het hergebruik van publieke overheidsdata [Likert, 1932]. Ook wordt gevraagd op een Likert-schaal van vijf punten aan te geven in hoeverre dit nu het geval is binnen de casus-organisatie. De resultaten van deze controlevragen kunnen worden gebruikt voor triangulatie van de interviewantwoorden, om daarmee de interne validiteit (paragraaf 3.4.2) te verbeteren.

4.1.6. Aanpak data-analyse

De analyse van de onderzoeksdata bestaat uit twee onderdelen: de analyse van de gehouden interviews en de analyse van de aanvullende documentatie.

Analyse interviews

De semi-gestructureerde interviews zijn zoals beschreven in paragraaf '3.2.1-Interviews' uitgevoerd. De analyse van de interviews is volgens de beschrijving in paragraaf '3.3-Gegevensanalyse' gedaan, rekening houdend met de overwegingen in paragraaf 3.4 over validiteit, betrouwbaarheid en ethische aspecten. Er is gevraagd om de invloed van verschillende factoren op de herbruikbaarheid van publieke overheidsdata aan te geven met behulp van een waarderingsschaal, om meer inzicht te krijgen in de huidige situatie en de ideale situatie. De data van de interviews is verwerkt in de onderzoeksresultaten, zoals beschreven in paragraaf 4.2.1.

Analyse aanvullende documentatie

Tijdens de semigestructureerde interviews is gevraagd om mogelijke ondersteunende documentatie. Deze documentatie is geanalyseerd zoals deze is beschreven in paragraaf '3.3-Gegevensanalyse'. De analyse van de aanvullende documentatie is verwerken in de onderzoeksresultaten, zie paragraaf '4.2.2-Resultaten documentanalyse'.

4.2. Resultaten

In de volgende paragrafen worden de resultaten van het empirische deel geanalyseerd. Eerst worden de resultaten van de semigestructureerde interviews geanalyseerd, daarna de aanvullende documentatie. Vervolgens worden deze resultaten per variabele toegelicht en wordt dit deel afgesloten met een samenvatting. In hoofdstuk '5-Discussie, conclusies en aanbevelingen', staan de conclusies en aanbevelingen die gebaseerd zijn op deze resultaten.

Op basis van de kwalitatieve, gevalideerde data is er een samenvatting gemaakt per indicator van variabelen van het conceptueel model. Deze samenvattingen zijn terug te vinden in paragraaf 'D.1-Samenvatting transcripties'. De analyse van de kwalitatieve resultaten van de controlevragen zijn weergegeven in paragraaf 'D.2-Resultaten triangulatie antwoorden'. In paragraaf 4.2.1 worden eerst de samenvatting per indicator van variabelen uit het conceptueel model weergegeven. Daarna, in paragraaf 4.2.2, worden de resultaten van de documentanalyse weergegeven, met als laatste een samenvatting van deze resultaten in paragraaf 4.2.3.

4.2.1. Interview resultaten per variabele in het onderzoeksmodel

Op basis van het conceptueel model van figuur '2.3-Conceptueel model' worden de resultaten van de interviews en de documentanalyse kort toegelicht.

Toepassen van Enterprise Architectuur principes en richtlijnen

Voor dit kernbegrip zijn er interview vragen gesteld over de principes en richtlijnen van EA en EA-BD, deze zijn hieronder samengevoegd. Alle respondenten kennen de concepten van EA en BD.

Respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker een
- Beleidsmedewerker twee
- Datamanager
- Product Owner DONL

Nemen van beslissingen (EA-1)

Managers maken nauwelijks gebruik van artefacten uit de EA om beslissingen te nemen, terwijl architecten hier juist wel gebruik van maken.

Coördinatie (EA-2)

Architecten op zowel ondernemingsniveau als binnen businessunits werken samen om de lokale en globale behoeften in balans te brengen. Er wordt gewerkt aan een gedeelde visie op de architectuur op zowel ondernemingsniveau als binnen businessunits, maar er kan nog verbetering worden aangebracht in de afstemming tussen deze twee niveaus.

Algemene acceptatie van EAM door de organisatie (EA-3)

De organisatie is op de hoogte van Enterprise Architecture Management (EAM), maar het gebruik ervan is nog heel beperkt. Er is wel al een begin gemaakt met de ontwikkeling van een federatief datastelsel binnen DO. De Data Catalog Vocabulary (DCAT) standaard zou kunnen worden gebruikt om de samenwerking tussen de verschillende EA's te verbeteren.

Algemene kenmerken van de organisatiecultuur (EA-4)

Binnen KOOP zijn er geen verschillende subculturen binnen het domein van EA. Er zijn duidelijke opdrachten en er is hulp beschikbaar voor EA als dat nodig is.

Culturele verschillen tussen architecten en managers (EA-5)

Er zijn geen grote verschillen in cultuur tussen architecten en managers binnen het domein van EA. De functie van enterprise architect wordt als een professie gezien en is daarmee een nuttige functie.

Culturele verschillen tussen verschillende architecten (EA-6)

Er kan weinig gezegd worden over verschillende subculturen tussen architecten binnen het domein van EA vanuit de perspectieven van DO en KOOP.

Functionele geschiktheid (EA-BD-1)

De functionele eisen zijn vaak wel duidelijk, de architectuur achter de data echter minder duidelijk. Er is een wettelijke basis die beschrijft dat data openbaar gemaakt moeten worden, zoals de Wet open overheid (Woo) en Wet hergebruik van overheidsinformatie (Who). Maar deze wetten geven weinig specifieke richtlijnen, waardoor de kwaliteit, volledigheid en integriteit van de data op data.overheid.nl een uitdaging blijft.

Betrouwbaarheid (EA-BD-2)

Er zijn geen specifieke regels voor beschikbaarheid en fouttolerantie voor data.overheid.nl, waardoor de betrouwbaarheid van de data beperkt is. Feedback van gebruikers kan helpen bij het verbeteren van datasets, door het toepassen van kwaliteitscriteria's en het garanderen van volledigheid, juistheid en nauwkeurigheid. Juristen zetten kanttekeningen bij het delen van data, omdat het moet voldoen aan de Uitvoeringswet Algemene Verordening Gegevensbescherming (AVG), maar ook aan de Wet openbaarheid van bestuur (Wob) en de Woo.

Prestatie-efficiëntie (EA-BD-3)

Op data.overheid.nl is de prestatie-efficiëntie beperkt, maar er wordt wel ruimte gegeven om dingen te doen. Er wordt onderzocht welke regels er opgesteld moeten worden om de prestaties te verbeteren. Het bijwerken van de data op data.overheid.nl gebeurt soms automatisch met behulp van Application Programming Interface (API)'s.

Bruikbaarheid (EA-BD-4)

Er is onvoldoende inzicht in hoe de data op data.overheid.nl worden gebruikt. Er wordt contact gezocht met gebruikers om succesverhalen te delen. Er komen ook dataverzoeken binnen, waarbij er tegen problemen worden aangelopen. Op politiek niveau is er een opdracht uit Europa om meer contact te krijgen met de gebruikers van de data. In de Tweede Kamer er is een nieuw bewindspersoon die hiermee bezig gaat.

Beveiliging (EA-BD-5)

Er zijn wetgevingseisen voor het openbaar maken van data. Er moet worden aangegeven waarom bepaalde informatie niet openbaar kan zijn, bijvoorbeeld vanwege privacybescherming of omdat het gevoelige informatie betreft. Het is niet altijd duidelijk wat wel of niet openbaar gemaakt mag worden, daarom wordt er vaak voor gekozen om twijfelgevallen niet openbaar te maken. Bij publieke overheidsdata zou in principe aangegeven moeten worden wat openbaar is en wat niet, maar dit gebeurt vaak niet. Data die op data.overheid.nl staat, zoals die van het Koninklijk Nederlands Meteorologisch Instituut (KNMI) mogen commercieel worden gebruikt. Door meer data te delen wordt de economische groei gestimuleerd.

Compatibility (EA-BD-6)

Het Nederlandse implementatie van het toepassingsprofiel 2.0 is gepland voor de zomer van 2022. Dit is een Europese standaard die de communicatie via metadata standaardiseert met behulp van de DCAT standaard van de World Wide Web Consortium (W3C). Er zijn verbeteringsinitiatieven aan de gang om data samen te voegen, maar de formaten en frequenties van updates lopen niet gelijk.

Algemeen Enterprise Architectuur en Big Data (EA-BD-1 t/m EA-BD-6)

Voor data.overheid.nl richt KOOP zich specifiek op het Europese toepassingsprofiel van DCAT en 90% van wat op data.overheid.nl staat voldoet aan deze standaard. Er kan afgeweken worden van deze richtlijnen, maar het belangrijkste uitgangspunt is om zoveel mogelijk de FAIR en DCAT richtlijnen te volgen. Er wordt geen onderscheid gemaakt tussen BD en data op data.overheid.nl. KOOP heeft zelf geen BD.

Toepassen van Data Governance principes en richtlijnen

Respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker een
- Beleidsmedewerker twee
- Datamanager

Visie (DG-BD-1)

De visie, is een interbestuurlijke data strategie voor huidige systeem en het opbouwen van een federatief datastelsel waarbij de data bij de bron blijven, maar wel centraal toegankelijk zijn. Als voorbeeld de tien basis

registraties van Nederland. Hiervoor zijn er door KOOP centrale functies beschikbaar gesteld die onder dit basisregistratiestelsel vallen. Er wordt onderzocht of dit uitgebreid kan worden met andere datasets en of de vindbaarheid en doorzoekbaarheid van data verbeterd kan worden, rekening houdend met de FAIR Data Principles. De visie op data is dat het zo efficiënt mogelijk verzameld en gebruikt wordt binnen de kaders van privacybescherming en wetgeving.

Leiderschap (DG-BD-2)

In Nederland is het bestuur extreem versnipperd. Er is de wens voor een centrale rol, die bij het BZK gelegd kan worden. Maar dit hoeft niet noodzakelijkerwijs op een hiërarchische manier op de hoogste plek binnen het ministerie te zijn. Leiderschap wordt steeds belangrijker, maar het landschap van open data is gedecentraliseerd. Het stimuleren van open data ligt meer bij de Open State Foundation (OSF).

Financiering Open data (DG-BD-3)

De financiering van data.overheid.nl wordt geregeld door BZK. Decentrale overheden en departementen zorgen ervoor dat data beschikbaar zijn en financieren dit zelf. BZK is verantwoordelijk voor het beheer van open data als het gaat over de Woo en de technische invulling ervan.

Zelf-ontwikkeld vermogen (DG-BD-4)

De Datamanager van KOOP analyseert welke databronnen beschikbaar zijn en ingezet kunnen worden, afhankelijk van de informatiebehoefte. Er is een wettelijke verplichting dat verzoeken voor hergebruik van data via data.overheid.nl kunnen worden gedaan. Het is de verantwoordelijkheid van de betreffende organisatie om deze verzoeken uit te voeren.

Communicatie (DG-BD-5)

De belangrijkste communicatie verloopt via data.overheid.nl, maar ook via Digitale Overheid en datacommunities.nl. KOOP speelt een belangrijke rol tussen de personen die informatie vragen en degenen die informatie aanbieden. Dit wordt voornamelijk verzorgd door de Datamanager.

Interoperabiliteit (DG-BD-6)

De wet Woo zal naar verwachting veel nieuwe data opleveren. Voor een open overheid is het belangrijk om duidelijk te maken wat de bruikbaarheid van deze data is. De overheid heeft de principes voor het gebruik van data opgesteld in samenwerking met de Organisation for Economic Co-operation and Development (OECD). Goede data is belangrijk en hierbij spelen de FAIR Data Principles een belangrijke rol. Door gebruik te maken van DCAT, spreekt iedereen dezelfde taal over data. De licentievoorraude voor deze data is open data en de FAIR Data Principles kunnen hierbij helpen.

Metadata (DG-BD-7)

Metadata kan beschouwd worden als het middelpunt van de wereld van data. Hoe beter en uniformer de metadata zijn, hoe gemakkelijker deze gedeeld kunnen worden. Het federatief datastelsel en de datacatalogus kunnen hierbij ondersteunend zijn. Momenteel is KOOP bezig met het herschrijven van het toepassingsprofiel van DCAT, met name voor metadata.

Toepassen van FAIR data management principes en richtlijnen

Alle respondenten kennen de concepten van FAIR Data en BD, maar het concept van FADAM was niet bekend bij de respondenten. Het concept FADAM is uitgelegd als een combinatie van de principes en richtlijnen van DG en FAIR Data.

Respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker een
- Beleidsmedewerker twee
- Datamanager
- Product Owner DONL

Management Commitment (FADAM-BD-1)

Er is nog veel stimulering nodig om data te delen. Vaak is het management zich niet bewust van de noodzaak hiervan. De wetgeving van de Woo speelt hierbij vaak een belangrijke rol, dan snapt het management het vaak wel waarom data delen belangrijk is. Door data openbaar te maken, wordt de herbruikbaarheid en effectiviteit van data verbeterd. Het is daarom belangrijk dat het management opdracht geeft om data openbaar te maken, je ben gehouden aan de wet. Dit kost geld en hiervoor zou een centraal budget beschikbaar gesteld kunnen worden. Het commitment van het management is belangrijk, want als het management het belangrijk vindt, wordt dit ook op de werkvloer opgepakt. De combinatie van commitment van het management en beschikbaarheid van budget is cruciaal. Overheden, gemeenten en departementen beschikken over heel veel data, waarvan een deel open data, maak ook gesloten data. De casus-organisatie gaat hier mogelijk later nog naar kijken.

Managementplanning (FADAM-BD-2)

Bij het management van KOOP wordt weinig aandacht besteed aan de verbetering van de kwaliteit van data, qua managementplanning. Binnen BZK wordt een planning voor vijf jaar gemaakt, zoals het inrichten van flexibele ontwikkelteams en het opzetten van een federatief datastelsel. Als de open data van onvoldoende kwaliteit is, kan dit leiden tot verkeerde conclusies met mogelijk gevaarlijke gevolgen. Het commitment van het management is daarom van groter belang dan alleen de managementplanning.

Vindbaarheid (FADAM-BD-3)

Nederland scoort hoog op het gebied van data-toegankelijkheid. Dit is een belangrijk aspect van data.overheid.nl en er is veel geïnvesteerd in het verbeteren hiervan. De basisregistraties zijn gemakkelijk te vinden door de stelselcatalogus, mede dankzij de twaalf eisen die hierop van toepassing zijn. Een belangrijk aspect hiervan is dat elk item een Uniform Resource Identifier (URI) heeft en unieke identifiers. Een goede vindbaarheid van metadata is net zo belangrijk, zodat er onderscheid kan worden gemaakt tussen het thema, doel, bestandsformaat en frequentie van bijwerkingen van een dataset.

Toegankelijkheid (FADAM-BD-4)

Er wordt gewerkt met de standaarden van W3C en Resource Description Framework (RDF) om metadata te beschrijven en API's beschikbaar te maken. De DCAT standaard wordt gebruikt om datasets op data.overheid.nl te beschrijven, zoals het bestandsformaat. Er is nog veel werk te doen om de herbruikbaarheid van de data laagdrempelig te maken, zodat niet alleen experts hiermee kunnen werken. Dit kan worden bereikt door de data op verschillende manieren beschikbaar te stellen.

Interoperabiliteit (FADAM-BD-5)

Er is veel aandacht voor het afstemmen van standaarden, omdat datasets gemakkelijker te combineren zijn als ze toegankelijk zijn. Er is geen wettelijk basis hiervoor, maar organisaties weten elkaar goed te vinden. Een aanbeveling is om een standaardisatie forum te organiseren waarin organisaties worden samengebracht. Hiervoor is een centrale coördinatie wenselijk en een lijst met standaarden.

Herbruikbaarheid (FADAM-BD-6)

Als alleen de vindbaarheid van de data is gegarandeerd, is de herbruikbaarheid alsnog beperkt. Als daarnaast de toegankelijkheid en interoperabiliteit worden verbeterd, neemt de herbruikbaarheid alleen maar verder toe. Door gebruik te maken van standaardisatie van dataformaten, zoals RDF, JavaScript Object Notation (JSON) en *Terse RDF Triple Language (Turtle)*, neemt de herbruikbaarheid van open data toe. De herbruikbaarheid van open data kan verder worden vergroot door die semantisch interoperabel te maken door bestaande standaarden en begripdefinities te gebruiken van de datacatalogus. Er zijn nog veel data die ontsloten moeten worden en dit heeft hoge prioriteit. De Nederlandse overheid heeft gekozen voor de Europese DCAT standaard voor de metadata, hiermee kan alles wat op data.overheid.nl staat eenvoudig worden overgezet naar de Europese open data site.

Ontwikkelen van Best Practices Community Models

Alle respondenten kennen het concept BD, maar het concept van CM was niet bekend bij alle respondenten. Het concept CM is uitgelegd als een set van *Best Practices* voor datacommunities.

Respondenten:

- Adviseur Content Standaarden
- Community manager
- Product Owner DONL

Samenwerking (CM-BD-1)

KOOP is in gesprek met veel partijen, waaronder een groep van vijf datacommunities zoals die van energie en onderwijs. Deze datacommunities hebben onderling geen contact met elkaar. Iedereen kan een datacommunity starten en mensen kunnen daar vrijblijvend aansluiten. Daar kunnen vragen gesteld en beantwoord worden en informatie gedeeld. De meeste communities zijn ontstaan tijdens de Corona-periode, waardoor er een behoefte was aan dit soort platformen. Nu er weer congressen worden georganiseerd, neemt deze behoefte misschien wel weer af.

Vaardigheden en training (CM-BD-2)

Er zijn geen specifieke trainingen nodig, het belangrijkste is een open houding en inhoudelijke kennis. De datacommunities.nl zijn laagdrempelig en is een uitleg video beschikbaar over data.overheid.nl. Er is een nieuwe Community manager aangesteld om mensen te ondersteunen bij de datacommunities, bijvoorbeeld door een roadshow. De grootste gebruikers van het systeem zijn voornamelijk technische mensen die er vertrouwd mee zijn, maar dat geldt niet voor nieuwkomers.

Openheid (CM-BD-3)

In het verleden was er veel angst om data te delen, maar dit begint te veranderen. Er wordt nu gekeken naar hoe data op een verantwoorde manier gedeeld kunnen worden. KOOP erkent het belang van de Wob, maar houdt ook rekening met de kosten die hiermee gepaard gaan. Ook de Woo draagt bij aan het openbaar maken van data, waarbij alles wat als open data wordt beschouwd, publiek toegankelijk en herbruikbaar moet zijn en gemakkelijk te vinden. Echter, de data binnen de datacommunities wordt vaak afgeschermd, wat een drempel kan vormen.

Technische infrastructuur (CM-BD-4)

Er is technische infrastructuur beschikbaar op data.overheid.nl, maar het uitgangspunt is om de data bij de bron te houden. Om data uit documenten beschikbaar te maken, kan dat via het Platform Open Overheidsinformatie (PLOOI) gedaan worden. Voor datacommunities.nl wordt de tool *Huddle* gebruikt, maar data.overheid.nl kan ook als een tool gezien worden.

Gemeenschappelijke participatie (CM-BD-5)

Er worden regelmatig bijeenkomsten en seminars georganiseerd waar gebruikers en stakeholders kennis en ervaringen kunnen delen. Binnen datacommunities zijn er ambassadeurs die het persoonlijk belangrijk vinden dat deze community bestaat en ervoor zorgen dat vragen snel beantwoord worden. Om te achterhalen wat de ambassadeurs van het data.overheid.nl platform vinden, kan men de ambassadeurs bij elkaar brengen, maar in de praktijk komt gemeenschappelijke participatie nog niet vaak voor. Het primaire doel van data.overheid.nl is het verkrijgen en verstrekken van data.

Economisch en bedrijfsmodel (CM-BD-6)

De open data op data.overheid.nl zijn openbaar en deze mogen gebruikt worden voor economische doelen. Er is echter geen ondersteuning voor economische en bedrijfsmodellen vanuit KOOP.

Juridische en ethische kwesties (CM-BD-7)

Er is veel aandacht voor bescherming van persoonsgegevens en de AVG moet goed in de gaten worden gehouden. Dit geldt ook voor verzoeken om informatie die KOOP gaat publiceren volgens de Wob en Woo. Tot nu toe zijn er geen ethische kwesties geweest, omdat het om openbare informatie gaat. Binnen de data communities zijn regels opgesteld waaraan men moet voldoen, anders kan de toegang worden ontzegd. Binnen data communities kunnen juristen actief zijn die invloed uitoefenen op betrokken partijen, zoals het delen van data voor de energietransitie.

Academische cultuur (CM-BD-8)

De Universiteit van Amsterdam (UvA) was de initiator van het platform Linked Data (LD) en volgt het nu op afstand. Het platform is nu ondergebracht bij een commerciële partij. De meeste deelnemers aan het platform zijn hoog opgeleid, wat leidt tot een behoorlijke wetenschappelijke basis voor het platform. Er komen ook dataverzoeken binnen op data.overheid.nl vanuit de wetenschappelijke wereld, maar dit is slechts een kleine groep.

Participatie (CM-BD-9)

Er zijn veel *ad hoc* overleggen, vooral op operationele niveau, maar er is weinig overleg op het bestuurlijk niveau. Een vorm van participatie zou kunnen zijn om de ambassadeurs en trekkers van elke data community bij elkaar te brengen om verbeteringen aan te brengen in data.overheid.nl, bijvoorbeeld door het versturen van nieuwsbrieven.

Engagement (CM-BD-10)

Dit gaat met name over data.overheid.nl voor de data communities en tracht actuele onderwerpen op de agenda te krijgen. Er zijn plannen voor een marketingstrategie met als doel mensen meer te betrekken bij data.overheid.nl en de data communities. Het geven van een kleine cursus of opleiding over hoe men data op data.overheid.nl kan gebruiken kan belanghebbenden helpen. Het doel is dat data.overheid.nl een platform wordt voor alle data, waarbij een goede zoekfunctie belangrijk is. Sociale media spelen ook een grote rol, waarmee onderwerpen over data communities onder de aandacht kunnen worden gebracht, bijvoorbeeld door iets te posten op LinkedIn.

Stimulansen (CM-BD-11)

De data communities bij data.overheid.nl worden beleidsmatig gestimuleerd. Samenwerking ontstaat vaak van onderaf, maar gebeurt nog niet vaak. Hier ligt wellicht een kans voor verbetering. De personen binnen de data communities hebben vaak te weinig tijd voor extra activiteiten. De Woo zal een grote aanjager zijn om data openbaar te maken.

Community Governance Models (CM-BD-12)

KOOP maakt geen gebruik van communities governance modellen voor data.overheid.nl, vaak wordt er nog in de waan van de dag geleefd.

4.2.2. Resultaten documentanalyse

In de interviews is meerdere keren verwezen naar externe documentatie, zoals wetgeving en gebruikte standaarden. Deze documenten zijn geanalyseerd en de resultaten hiervan zijn beschreven in tabel '4.2-Samenvatting documentanalyse'. De details hiervan zijn weergegeven in bijlage D, paragraaf D.3. Geen van deze documenten heeft betrekking op de onderzoeksvragen of het conceptueel model en daarom zijn ze niet verder meegenomen in dit onderzoek. De casus-organisatie gebruikt deze documenten als wettelijke basis voor het opstellen van beleid.

Naam / Locatie	Analyse
Wet openbaarheid van bestuur https://wetten.overheid.nl/BWBR0005252/2018-07-28	Het betreft wetgeving waar aan voldaan moest worden en is inmiddels vervangen door de Woo. Dit is relevant, geldig en toepasbaar voor de huidige situatie van de casus-organisatie. Voor aanbevelingen uit dit onderzoek en verdere onderzoeken is het echter niet relevant, wel geldig, maar niet toepasbaar. Zie tabel 'D.2-Analyse document Wet openbaarheid van bestuur' voor de details.
Wet open overheid https://wetten.overheid.nl/BWBR0045754/2022-08-01	Het betreft wetgeving waar nu aan voldaan moet worden, ter vervanging van de Wob. Dit is relevant, geldig en toepasbaar voor de huidige situatie van de casus-organisatie. Voor aanbevelingen uit dit onderzoek en verdere onderzoeken is het echter niet relevant, wel geldig, maar niet toepasbaar. Zie tabel 'D.3-Analyse document Wet open overheid' voor de details.
Wet hergebruik van overheidsinformatie https://wetten.overheid.nl/BWBR0036795/2021-07-01	Het betreft wetgeving waar aan voldaan moet worden. Dit is relevant, geldig en toepasbaar voor de huidige situatie van de casus-organisatie. Voor aanbevelingen uit dit onderzoek en verdere onderzoeken is het echter niet relevant, wel geldig, maar niet toepasbaar. Zie tabel 'D.4-Analyse document Wet hergebruik van overheidsinformatie' voor de details.
Data Catalog Vocabulary https://www.w3.org/TR/vocab-dcat-2/	Dit is een internationale standaard die wordt gebruikt binnen de casus-organisatie. Dit is relevant, geldig en toepasbaar voor de huidige situatie van de casus-organisatie. Voor aanbevelingen uit dit onderzoek en verdere onderzoeken is het echter niet relevant, wel geldig, maar niet toepasbaar. Zie tabel 'D.5-Analyse document Data Catalog Vocabulary' voor de details.

Tabel 4.2: Samenvatting documentanalyse

4.2.3. Samenvatting van de resultaten

Alle interviewvragen zijn beantwoord door minimaal drie verschillende respondenten, zoals aangegeven in figuur '4.1-Respondenten en onderzoeksvragen', samen met de resultaten van de controlevragen van deze respondenten. Dit geeft een betrouwbaar beeld van de casus-organisatie op het gebied van de dimensies van het conceptueel model in figuur '2.3-Conceptueel model'. De samenvattingen bevatten de kwalitatieve data van meerdere respondenten, waardoor de constructvaliditeit verzekerd is.

Resultaat interviews

De resultaten van de data van de interviews per kernbegrip staan hieronder beschreven.

Enterprise Architectuur

Volgens de respondenten die de vragen over EA en EA-BD principes en richtlijnen hebben beantwoord, worden deze in de huidige situatie beperkt toegepast. Dit blijkt uit de analyse van de kernbegrippen, zoals weergegeven in tabel 'D.1-Analyse kernbegrippen'. Dit is in lijn met de indicatoren van EA en EA-BD, zoals eerder beschreven. Er zijn bij een aantal indicatoren verbetertrajecten in gang gezet of aangegeven waar verbeteringen mogelijk zijn. Daarnaast vinden de respondenten dat de EA en EA-BD principes en richtlijnen "toepasbaar" zijn om de herbruikbaarheid van publieke overheidsdata te kunnen verbeteren. Twee respondenten gaven aan dat dit "randvoorwaardelijk" is, zoals weergegeven in tabel D.1.

Data Governance

Volgens de respondenten die de vragen over de principes en richtlijnen van DG-BD hebben beantwoord, worden deze in de huidige situatie toegepast, zoals weergegeven in tabel 'D.1-Analyse kernbegrippen'. Dit is in lijn met de indicatoren van DG-BD zoals hierboven beschreven. Er zijn verbetertrajecten in gang gezet voor een aantal indicatoren, er is een duidelijke visie op basis van wet- en regelgeving en er wordt gebruik gemaakt van standaarden. Daarnaast vinden deze respondenten dat de principes en richtlijnen van DG en BD "goed toepasbaar" zijn om de herbruikbaarheid van publieke overheidsdata te verbeteren, waarbij één respondent geeft aangegeven dat dit "noodzakelijk" is, zoals weergegeven in tabel D.1.

FAIR Data Management

Volgens de respondenten die de vragen over FADAM–BD principes en richtlijnen hebben beantwoord, worden deze in de huidige situatie nauwelijks toegepast zoals weergegeven in tabel 'D.1-Analyse kernbegrippen'. Voor sommige indicatoren zijn er duidelijk acties ondernomen, maar voor andere indicatoren wordt nog niet die aandacht gegeven die misschien nodig is, met name voor management commitment en management-planning. Dit is in lijn met de indicatoren van FADAM–BD zoals hierboven beschreven. Daarnaast geven deze respondenten aan dat de FADAM en BD principes en richtlijnen "noodzakelijk" zijn om de herbruikbaarheid van publieke overheidsdata te kunnen verbeteren. Eén respondent geeft aan dat dit "randvoorwaardelijk" is, tabel D.1.

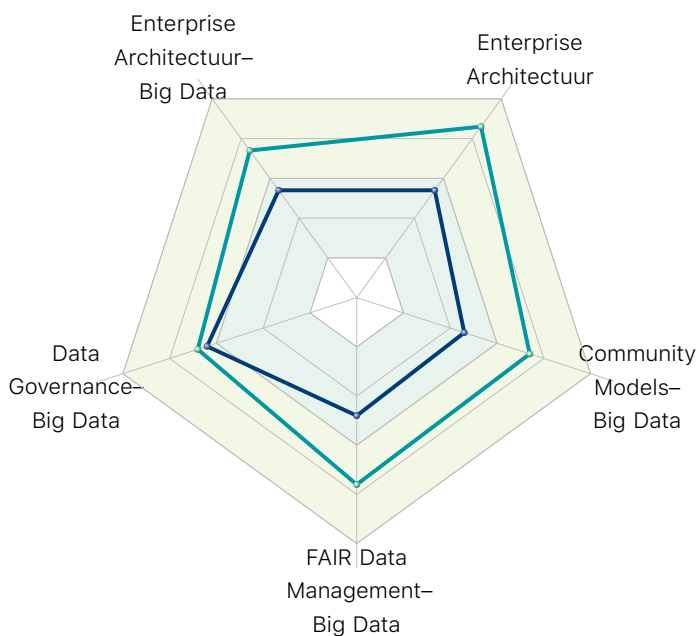
Community Models

Volgens de antwoorden van de respondenten op de vragen over *Best Practices* voor CM–BD, worden deze in de huidige situatie nauwelijks toegepast zoals weergegeven in tabel 'D.1-Analyse kernbegrippen'. Er worden acties ondernomen voor enkele indicatoren en voor andere indicatoren wordt aangegeven dat er verbeteringen mogelijk zijn. Dit is in lijn met de indicatoren van CM–BD zoals hierboven beschreven. De respondenten vinden dat de *Best Practices* voor CM–BD "verplicht" zijn om de herbruikbaarheid van publieke overheidsdata te verbeteren. Eén respondent geeft aan dat dit "randvoorwaardelijk" is, zoals weergegeven in tabel D.1. Deze *Best Practices* bevinden vooral op een operationeel niveau in plaats van op een beleidsniveau.

Resultaat kwalitatief onderzoek

De resultaten van de controlevragen zijn weergegeven in figuur '4.2-Analyse kernbegrippen'. Hierin is de achtergrond van het figuur onderverdeeld in een blauw en groen gedeelte. Het blauwe deel (binnenste) geeft aan dat de waarde kleiner is dan nul, het groen gedeelte (buitenste) geeft aan dat de waarde groter is dan nul, zoals beschreven in paragraaf 'D.2.1-Analyse kernbegrippen en dimensies'.

De blauwe lijn (binnenste) geeft de huidige situatie weer zoals aangegeven door de respondenten. De groene lijn (buitenste) geeft de bruikbaarheid aan van een dimensie waarmee de herbruikbaarheid van publieke overheidsdata kan worden verbeterd. Dit geldt voor alle dimensies, waarbij EA en FA-



Figuur 4.2: Analyse kernbegrippen

DAM het belangrijkste zijn, maar ook de CM worden als belangrijk aangegeven.

Een opmerking hierbij is dat sommige van de geïnterviewde respondenten niet langer bij de betreffende afdeling van de casus-organisatie werken. De interviewdata zijn gevalideerd door de respondenten op het moment dat ze nog werkzaam waren bij de betreffende afdeling, zoals aangegeven in tabel '4.1-Respondenten en onderzoeksvragen'. Dit heeft geen invloed op de kwaliteit van de data in dit onderzoek, maar kan wel invloed hebben op de toepasbaarheid van de resultaten van dit onderzoek in de casus-organisatie.

De conclusies en aanbevelingen van dit empirisch onderzoek zijn weergegeven in hoofdstuk '5-Discussie, conclusies en aanbevelingen'.

5

Discussie, conclusies en aanbevelingen

In dit hoofdstuk worden de resultaten van het onderzoek besproken in paragraaf '5.1-Discussie - reflectie'. De conclusie van het onderzoek is te vinden in paragraaf '5.2-Conclusies'. Het hoofdstuk wordt afgesloten met de aanbevelingen voor de praktijk en verder wetenschappelijk onderzoek, beschreven in respectievelijk paragraaf 5.3 en paragraaf 5.4.

5.1. Discussie - reflectie

We kunnen de onderzoeksvragen beantwoorden door het conceptueel model van figuur '2.3-Conceptueel model' en de resultaten van het empirisch onderzoek uit hoofdstuk 4 te bekijken. Er zijn meerdere respondenten geïnterviewd en bevestigd over de kernbegrippen van het conceptuele model. Dit gebeurde door middel van zowel interviewvragen over de verschillende onderwerpen van de verschillende dimensies, als door controlevragen op een kwalitatieve waarderingsschaal. De resultaten van het empirisch onderzoek zijn per kernbegrip samengevat in paragraaf 4.2.3. Samen met de resultaten en conclusies van uit het theoretisch kader in paragraaf 2.3 worden ze besproken.

Er zijn geen opvallende verschillen tussen deze resultaten uit de interviews en de controlevragen waargenomen. Echter, omdat het empirisch onderzoek met een beperkt aantal respondenten is uitgevoerd, kunnen er geen volledig gevalideerde conclusies worden getrokken. Binnen de casus-organisatie zijn er geen andere functionarissen in het werkveld waarin dit onderzoek is uitgevoerd, waardoor de constructvaliditeit niet verder versterkt kan worden. Aangezien het aantal functionarissen in dit werkveld beperkt is, is het belangrijk dat er voldoende tijd wordt uitgetrokken voor de interviews.

Hieronder volgen de conclusies per dimensie van het conceptueel model uit paragraaf 2.3.7.

5.1.1. Toepassen van Enterprise Architectuur principes en richtlijnen

De analyse was opgesplitst in twee dimensies uit de wetenschap: de principes en richtlijnen van Enterprise Architectuur (EA) en de principes en richtlijnen van EA-Big Data (BD). Als we kijken naar figuur '4.2-Analyse kernbegrippen', worden deze twee dimensies gelijkwaardig beoordeeld en dit komt overeen met wat uit de interviewvragen is gebleken en met de huidige situatie binnen de casus-organisatie. Het is duidelijk dat principes en richtlijnen van zowel EA als EA-BD zeer belangrijk worden gevonden binnen de casus-organisatie voor het verbeteren van de herbruikbaarheid van publieke overheidsdata. Er kunnen stappen worden gezet op het gebied van besluitvorming en het gebruik van EA artefacten, coördinatie en Enterprise Architecture Management (EAM). Het doel van de casus-organisatie is om de herbruikbaarheid van data op data.overheid.nl te verbeteren, in lijn met de Wet open overheid (Woo) en aankomende Europese regelgeving. EA wordt als goed toepasbaar beschouwd door de casus-organisatie om de herbruikbaarheid van publieke overheidsdata te verbeteren, maar wordt momenteel nog beperkt toegepast, dus is het belangrijk om dit verder te ontwikkelen.

5.1.2. Toepassen van Data Governance principes en richtlijnen

Er is onderzoek gedaan naar de principes en richtlijnen van Data Governance (DG)–BD uit de wetenschap. Bij de casus-organisatie worden deze principes en richtlijnen toegepast bij het maken en invoeren van beleid om de herbruikbaarheid van publieke overheidsdata te verbeteren. Dit wordt beschouwd als belangrijke voorwaarde en de principes en richtlijnen zijn goed toepasbaar. Waar mogelijk wordt er gewerkt aan het instellen van een centrale rol op het gebied van leiderschap en een duidelijke visie op data en DG. Metadata worden beschouwd als het belangrijkste middel om de herbruikbaarheid van publieke overheidsdata te verbeteren. Hierin zijn de eerste stappen gezet binnen de casus-organisatie, het is belangrijk om dit verder uit te bouwen.

5.1.3. Toepassen van FAIR Data Management principes en richtlijnen

Er is onderzocht gedaan naar hoe de principes en richtlijnen van FAIR Data Management (FADAM)–BD uit de wetenschap worden toegepast in de casus-organisatie. De toepassing van de FAIR Data Principles krijgt steeds meer aandacht in de casus-organisatie, maar is nog beperkt. De principes en richtlijnen worden als noodzakelijk beschouwd, vooral op het gebied van management commitment van FADAM. Er zijn nog stappen nodig om de herbruikbaarheid van publieke overheidsdata te verbeteren door toepassing van de FAIR principes en richtlijnen. Een goede managementplanning wordt beschouwd als belangrijk voor het verbeteren van de kwaliteit van de data. FADAM wordt in de casus-organisatie beschouwd als noodzakelijk om de herbruikbaarheid van publieke overheidsdata te verbeteren, dus is het belangrijk om dit verder te ontwikkelen.

5.1.4. Ontwikkelen van Best Practices Community Models

Er is onderzocht gedaan naar de manier waarop de *Best Practices* van Community Models (CM)–BD worden toegepast binnen de casus-organisatie. Deze *Best Practices* worden beperkt nageleefd en bevinden zich voornamelijk op het operationele niveau. De Corona-pandemie heeft aangetoond dat het belangrijk kan zijn om deze *Best Practices* op te volgen om de herbruikbaarheid van publieke overheidsdata te verbeteren. Er zijn initiatieven genomen op het gebied van samenwerking, openheid, gemeenschappelijke participatie, engagement en stimulansen. Door meer aandacht aan deze aspecten te besteden, heeft dit een duidelijke positieve invloed op de herbruikbaarheid van publieke overheidsdata. Dit wordt binnen de casus-organisatie beschouwd als verplichting om de herbruikbaarheid van publieke overheidsdata te verbeteren, dus is het belangrijk om hier verder mee aan de slag te gaan.

5.1.5. Sturing op beleid van publieke overheidsdata

Als we het conceptueel model van dit onderzoek bekijken, afgebeeld in figuur 2.3, kunnen we zien dat het toepassen van principes en richtlijnen van FADAM, DG en EA een directie invloed heeft op de sturing van beleid voor publieke overheidsdata. In de casus-organisatie worden deze principes en richtlijnen ook gebruikt om de sturing van het beleid rond hergebruik van publieke overheidsdata te verbeteren. Door de principes en richtlijnen van FADAM, BD en EA verder te ontwikkelen in deze richting, kan de sturing van het beleid verbeterd worden en daarmee ook de herbruikbaarheid van publieke overheidsdata. Dit resultaat is in lijn met het adviesrapport van de Koninklijke Nederlandse Academie van Wetenschappen (KNAW) waarin aangegeven is dat de toegankelijkheid en herbruikbaarheid van publieke overheidsdata verbeterd kan worden door onder andere het toepassen van FAIR Data Principles [knaW, 2018]. Deze afhankelijke variabele van het conceptueel model is daarmee gevalideerd binnen de casus-organisatie.

5.1.6. Herbruikbaarheid van publieke overheidsdata verbeteren

Wanneer we het conceptueel model van dit onderzoek bekijken, kunnen we zien dat de *Best Practices* voor CM een modererende werking hebben op de sturing op beleid van publieke overheidsdata en het verbeteren van de herbruikbaarheid van publieke overheidsdata. In de casus-organisatie worden deze *Best Practices* niet gebruikt om de sturing van het beleid rond hergebruik van publieke overheidsdata te verbeteren, maar ze hebben wel een duidelijke positieve invloed op de herbruikbaarheid van data. Binnen de casus-organisatie wordt hier wel aandacht aan besteed, maar dit is nog beperkt.

Naast de sturing op het beleid voor publieke overheidsdata is het ook belangrijk om *Best Practices* te ontwikkelen op het gebied van CM. Toch blijkt duidelijk dat deze *Best Practices* belangrijk zijn en een positieve invloed hebben op de herbruikbaarheid van publieke overheidsdata. Hierdoor is het conceptueel model van figuur 2.3 toepasbaar en in de praktijk gevalideerd, en kunnen er conclusies en aanbevelingen gedaan worden op basis hiervan.

5.1.7. Reflectie onderzoeksproces

Dit onderzoek was een enkelvoudige casusstudie op beleidsniveau binnen de casus-organisatie en had een beperkt aantal functionarissen als mogelijke respondenten. Als we dit onderzoek vergelijken met figuur '2.1-Vereenvoudigde conceptueel model van EAM voor gedecentraliseerde ondernemingen', is het uitgevoerd op "Enterprise Level" specifiek op het niveau van de "Enterprise Level Manager" en "Coördinatie". Binnen de casus-organisatie wordt het beleid voor publieke overheidsdata gemaakt door Beleidsmedewerkers van directie Digitale Overheid (DO), terwijl de coördinatie tussen "Enterprise level" en "Business unit level" wordt verzorgd door functionarissen van Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP). Hierdoor is er niet gekeken naar de uitvoering van dit beleid op "Business Unit" niveau wat invloed heeft op de kwaliteitsaspecten van dit onderzoek.

De kwaliteitsaspecten van dit onderzoek zijn beschreven in paragraaf 3.4. Voor dit onderzoek zijn meerdere menselijke bronnen gebruikt en er is een vorm van triangulatie toegepast door middel van controlevragen, waardoor de constructvaliditeit gewaarborgd is. Om de interne validiteit zo veel mogelijk te waarborgen, zijn de beschreven technieken gebruikt en is geprobeerd verklaringen te geven op basis van de theorie om bias te voorkomen. De interviews zijn geanonimiseerd en samengevat, waardoor de antwoorden van dit onderzoek minder controleerbaar zijn en de interne validiteit beperkt is. De externe validiteit is ook beperkt, omdat dit onderzoek alleen is uitgevoerd binnen Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK). Hierbij aangegeven dat dit de enige plek is in Nederland waar onderzoek op dit niveau kan worden uitgevoerd. Dit onderzoek is uitgevoerd volgens de normen van "De Nederlandse Gedragscode Wetenschapsbeoefening" om de betrouwbaarheid te waarborgen.

Met betrekking tot de ethische aspecten is dit onderzoek gebaseerd op vrijwillige deelname en is de privacy van de respondenten gewaarborgd door de data te samenvatten zodat deze niet herleidbaar zijn naar een specifieke functionaris. De interviewdata is getranscribeerd en gevalideerd door de respondenten. Vervolgens zijn de beeld- en geluidsopnames verwijderd en zijn alleen de geanonimiseerde, getranscribeerde data gebruikt. Vanwege de publieke verantwoordelijkheid van BZK streeft men naar openheid en transparantie en wordt dit onderzoek openbaar gedeeld via de Open Universiteit (OU), met de vereiste schriftelijke toestemming.

In paragraaf '1.3-Probleemstelling' is vermeld dat de term publieke overheidsdata is gekozen in plaats van specifiek te verwijzen naar open data en semi-open data. Dit komt doordat de definitie van open data duidelijk is, terwijl er geen definitie bestaat voor semi-open data. Voor dit onderzoek is daarom een eigen definitie aan semi-open data opgesteld in bijlage A, paragraaf 'A.1.1-Publieke overheidsdata', wat resulteerde in figuur '1.2-Publieke overheidsdata'. Het verschil tussen deze twee kan klein of substantieel zijn, en er ook een juridisch aspect verbonden aan de openheid van data. Daarom is gekozen voor de algemene term publieke overheidsdata. Dit blijkt ook uit de FAIR Data Principles in tabel 'A.4-FAIR leidende principes', waarin geen onderscheid wordt gemaakt tussen open of gesloten data, maar waarin, indien nodig, het protocol authenticatie- en autorisatieprocedure moet ondersteunen. Op basis hiervan zijn de conclusies en aanbevelingen van dit onderzoek opgesteld.

5.2. Conclusies

De hoofdvraag van dit onderzoek is:

Kan het toepassen van principes van FAIR Data Management aangevuld met algemene principes en richtlijnen, rond Enterprise Architectuur en Data Governance enerzijds en Community Models Best Practices anderzijds, helpen om de sturing op het beleid rond hergebruik van publieke overheidsdata te verbeteren?

Er is bij het BZK een kwalitatief onderzoek uitgevoerd om antwoord te vinden op deze vraag. De resultaten van dit onderzoek tonen aan dat de principes en richtlijnen van FADAM kunnen helpen bij het verbeteren van de sturing op beleid voor hergebruik van publieke overheidsdata. Bovendien blijkt uit de wetenschappelijke literatuur dat de algemene principes en richtlijnen van DG en EA een positieve invloed hebben op de sturing op beleid binnen de casus-organisatie. Door het verbeteren van de sturing op beleid zal de herbruikbaarheid van publieke overheidsdata toenemen en dit wordt verder versterkt door het ontwikkelen van *Best Practices* van CM.

Dit onderzoek laat zien dat door de principes en richtlijnen van FAIR Data Management te volgen, de herbruikbaarheid van publieke overheidsdata verbeterd kan worden. Het management-deel van FADAM richt zich op commitment en planning, terwijl FAIR Data zich richt op het verbeteren van vindbaarheid, toegankelijkheid, interoperabiliteit en herbruikbaarheid van data. Door de combinatie van algemene principes en richtlijnen om de sturing op het beleid te verbeteren, en tegelijkertijd de vindbaarheid, toegankelijkheid, interoperabiliteit en herbruikbaarheid van data te verbeteren, is het mogelijk om FADAM toe te passen voor het verbeteren van de herbruikbaarheid van publieke overheidsdata. Deze herbruikbaarheid wordt versterkt door het ontwikkelen van *Best Practices* van CM.

5.3. Aanbevelingen voor de praktijk

De conclusie van dit onderzoek is dat principes en richtlijnen van FADAM kunnen helpen aan het verbeteren van de herbruikbaarheid van publieke overheidsdata. Het is aan te bevelen om deze principes en richtlijnen van FADAM verder te ontwikkelen binnen de Nederlandse overheid om doelstelling van hergebruik te bevorderen en de publieke overheidsdata zo toegankelijk mogelijk te maken voor burgers, bedrijven en instellingen. Daarbij is het belangrijk om de principes en richtlijnen van DG en EA in acht te nemen, aangezien uit wetenschappelijk onderzoek is gebleken dat deze ondersteuning bieden voor de sturing op beleid. Bovendien is gebleken dat de toepassing van *Best Practices* van CM een positieve invloed heeft op het hergebruik van publieke overheidsdata. Daarom verdient het de aanbeveling om deze *Best Practices* van CM te implementeren binnen de Nederlandse overheid. Met de Nederlandse overheid bedoelen we zowel de rijksoverheid als provincies en gemeenten.

Voor de casus-organisatie is het aan te bevelen om te richten op de metadata van de publieke overheidsdata. Dit komt voort uit de principes en richtlijnen van DG, waarbij Master Data Management (MDM) een belangrijk aspect is [Khatri, 2010]. Bovendien speelt metadata ook een belangrijke rol in FAIR, waarbij de metadata vindbaarheid, toegankelijkheid, interoperabiliteit en herbruikbaarheid moet zijn [Wilkinson, 2016].

Management commitment en management planning zijn een belangrijke aspecten van de principes en richtlijnen van FADAM waarvan de toegevoegde waarde wordt onderschat [dFigueiredo, 2019]. Daarom is het aan te bevelen dat het management deze twee managementprincipes van FADAM ter harte neemt.

5.4. Aanbevelingen voor verder onderzoek

Er is wetenschappelijk onderzoek verricht naar de algemene principes van DG en EA en de vertaling hiervan naar beleidslijnen [Josey, 2016; OU, 2009; vdBerg, 2004]. Ook is er onderzoek gedaan naar specifieke architectuurprincipes voor BD en publieke overheidsdata [Henderson, 2017; Josey, 2016; OU, 2009]. In dit onderzoek hebben we geen verschil kunnen vinden tussen open data of semi-open data, behalve op juridische gebied. Deze aspecten maken ook deel uit van de FAIR Data Principles, wat mogelijk samen kan bijdragen aan

de theorievorming rond publieke overheidsdata. Er is echter nog weinig onderzoek gedaan naar CM, hoewel enkele *Best Practices* bekend zijn [Curry, 2010; Lyon, 2012b]. In dit onderzoek worden deze wetenschappelijke onderzoeken gecombineerd met de principes en richtlijnen van FAIR Data Principles en FADAM [dFigueiredo, 2019; Hasnain, 2018; Wilkinson, 2016].

Dit onderzoek is een enkelvoudige casusstudie op beleidsniveau binnen BZK. Om op de lange termijn resultaten te verkrijgen, is verder onderzoek nodig. Toekomstig onderzoek kan deze resultaten verifiëren en mogelijke conclusies trekken over de ontwikkeling van de herbruikbaarheid van publieke overheidsdata. Bovendien kan verder onderzoek op een lager niveau worden uitgevoerd, bijvoorbeeld specifiek gericht op de herbruikbaarheid van publieke overheidsdata van specifieke overheidsorganisaties, zoals uitvoeringsorganisaties, provincies en gemeentes. Dit vereist echter een andere benadering, omdat er veel overheidsorganisaties zijn die data publiceren.

Over het algemeen kan worden gesteld dat er nog veel mogelijkheden zijn voor verder onderzoek naar DG, FAIR Data Principles, FADAM, maar ook EA en CM, in relatie tot het verbeteren van de herbruikbaarheid van publieke overheidsdata. Het conceptueel model uit figuur '2.3-Conceptueel model' is gevalideerd in de casusorganisatie en kan dienen als basis voor verder empirisch of theoretisch onderzoek. De Nederlandse overheid is groot en de verschillen tussen organisaties kunnen daardoor groot zijn, wat veel mogelijkheden biedt voor verder onderzoek in combinatie met dit conceptueel model. Dit onderzoek is uitgevoerd op het beleidsniveau van de Nederlandse overheid, waarbij het BZK verantwoordelijk is.

Laat dit onderzoek als inspiratie dienen voor verder onderzoek om zo de herbruikbaarheid van publieke overheidsdata te verbeteren en deze data toegankelijker te maken voor burgers, bedrijven en instellingen.

Bibliografie

Boek

- Curry, E., Freitas, A. & O’Riáin, S. (2010). The role of community-driven data curation for enterprises. In *Linking enterprise data* (pp. 25–47). Springer.
- de Figueiredo, G. B., Moreira, J. L. R., de Faria Cordeiro, K. & Campos, M. L. M. (2019). Aligning dmbok and open government with the fair data principles. In G. Guizzardi, F. Gailly & R. Suzana Pitangueira Maciel (Red.), *Advances in conceptual modeling* (Hfdstk. Chapter 2, pp. 13–22). Lecture Notes in Computer Science. doi:[10.1007/978-3-030-34146-6_2](https://doi.org/10.1007/978-3-030-34146-6_2)
- Doorewaard, H. [Hans], Kil, A. & van de Ven, A. (2019). *Praktijkgericht kwalitatief onderzoek* (2de ed.). Boom uitgevers Amsterdam. Verkregen van <https://www.praktijkgerichtkwalitatiefonderzoek2edruk.nl>
- Henderson, D., Earley, S. & Sebastian-Coleman, L. (2017). *Dama-dmbok: Data management body of knowledge* (2de ed.). doi:[2017941854](https://doi.org/2017941854)
- Josey, A. (2016). *Togaf® version 9.1-a pocket guide*. Van Haren.
- Kepner, C. H. & Tregoe, B. B. (1997). *The new rational manager: An updated edition for a new world*. BookBaby.
- Lankhorst, M. (2009). *Enterprise architecture at work*. Springer.
- OU, O. U. N. (2009). *Informatie- en procesarchitectuur*. Heerlen: Open Universiteit.
- Ross, J. W., Weill, P. & Robertson, D. (2006). *Enterprise architecture as strategy: Creating a foundation for business execution*. Harvard business press.
- Saunders, M., Lewis, P. & Thornhill, A. (2019). *Research methods for business students* (Eighth). Harlow: Pearson Education Limited.
- van den Berg, M. & van Steenberg, M. (2004). *Dya: Stap voor stap naar professionele enterprise-architectuur*. TenHagenStam Uitgevers.
- Verschuren, P. J. M. & Doorewaard, H. [H.]. (2015). *Het ontwerpen van een onderzoek* (Vijfde druk). Amsterdam: Boom Lemma uitgevers.
- Wagter, R., Van Den Berg, M., Luijpers, J. & Van Steenberg, M. (2005). *Dynamic enterprise architecture: How to make it work*. John Wiley en Sons.

Academische Artikelen

- Avery, A. & Cheek, K. (2015). Analytics governance: Towards a definition and framework.
- Al-Badi, A., Tarhini, A. & Khan, A. I. (2018). Exploring big data governance frameworks. *Procedia Computer Science*, *141*, 271–277.
- Baijens, J., Helms, R. W. & Velstra, T. (2020). Towards a framework for data analytics governance mechanisms.
- Bommel, P. v., Buitenhuis, P., Hoppenbrouwers, S. & Proper, E. (2007). Architecture principles—a regulative perspective on enterprise architecture. *Enterprise modelling and information systems architectures—concepts and applications*.
- Bruckner, T., Voříšek, J., Buchalceková, A., Stanovská, I., Chlapek, D. & Řepa, V. (2012). Tvorba informačních systémů. *Praha, Česká republika: Grada Publishing*.
- Burmeister, F., Drews, P. & Schirmer, I. (2018). Towards an extended enterprise architecture meta-model for big data—a literature-based approach. *Americas Conference on Information Systems, 2018*(AMCIS 2018).
- Donker, F. W., Braggaa, R. & van Loenen, B. (2019). Hergebruikers van open data in beeld.
- Gong, Y. & Janssen, M. (2021). Roles and capabilities of enterprise architecture in big data analytics technology adoption and implementation. *Journal of theoretical and applied electronic commerce research*, *16*(1), 37–51.

- Janssen, M. & Kuk, G. (2016). Big and open linked data (bold) in research, policy, and practice. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 3–13.
- Jonkers, H., Lankhorst, M., Van Buuren, R., Hoppenbrouwers, S., Bonsangue, M. & Van Der Torre, L. (2004). Concepts for modeling enterprise architectures. *International Journal of Cooperative Information Systems*, 13(03), 257–287.
- Khatri, V. & Brown, C. V. (2010). Designing data governance. 53(1), 148–152.
- Kotusev, S. (2017). Conceptual model of enterprise architecture management. *International Journal of Cooperative Information Systems*, 26(03), 1–36. doi:10.1142/s0218843017300017
- Likert, R. (1932). A technique for the measurement of attitudes. *Archives of psychology*.
- Lnenicka, M. & Komarkova, J. (2019). Developing a government enterprise architecture framework to support the requirements of big and open linked data with the use of cloud computing. *International Journal of Information Management*, 46, 124–141.
- Mons, B., van Haagen, H., Chichester, C., den Dunnen, J. T., van Ommen, G., van Mulligen, E., ... Hammond, J. (2011). The value of data. *Nature genetics*, 43(4), 281–283.
- Otto, B. (2011b). Organizing data governance: Findings from the telecommunications industry and consequences for large service providers. *Communications of the Association for Information Systems*, 29(1), 3.
- Peffer, K., Tuunanen, T., Rothenberger, M. A. & Chatterjee, S. (2007). A design science research methodology for information systems research. *Journal of management information systems*, 24(3), 45–77.
- Rau, K. G. (2004). Effective governance of it: Design objectives, roles, and relationships. *Information Systems Management*, 21(4), 35–42.
- Sowa, J. F. & Zachman, J. A. [John A.]. (1992). Extending and formalizing the framework for information systems architecture. *IBM systems journal*, 31(3), 590–616.
- Welle Donker, F. & van Loenen, B. (2017). How to assess the success of the open data ecosystem? *International Journal of Digital Earth*, 10(3), 284–306.
- Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., ... Mons, B. (2016). The fair guiding principles for scientific data management and stewardship. *Sci Data*, 3(1), 160018. doi:10.1038/sdata.2016.18
- Zachman, J. A. [J. A.]. (1987). A framework for information systems architecture. *IBM Systems Journal*, 26(3), 276–292. doi:10.1147/sj.263.0276
- Zuiderwijk, A. & Janssen, M. (2014). Open data policies, their implementation and impact: A framework for comparison. *Government Information Quarterly*, 31(1), 17–29.

Conferenties

- Bargh, M. S., Choenni, S. & Meijer, R. (2016). Meeting open data halfway: On semi-open data paradigm. In *Proceedings of the 9th international conference on theory and practice of electronic governance* (pp. 199–206).
- Dedić, N. & Stanier, C. (2016). Towards differentiating business intelligence, big data, data analytics and knowledge discovery. In *International conference on enterprise resource planning systems* (pp. 114–122). Springer.
- Hasnain, A. & Rebholz-Schuhmann, D. (2018). Assessing fair data principles against the 5-star open data principles. In *European semantic web conference* (pp. 469–477). Springer.
- Janssen, M., Matheus, R. & Zuiderwijk, A. (2015). Big and open linked data (bold) to create smart cities and citizens: Insights from smart energy and mobility cases. In *International conference on electronic government* (pp. 79–90). Springer.
- Lyon, L., Ball, A. [Alexander], Duke, M. & Day, M. (2012). Developing a community capability model framework for data-intensive research. In *Proceedings of the 9th international conference on the preservation of digital objects* (pp. 9–16). Citeseer.
- Otto, B. (2011a). A morphology of the organisation of data governance. In *Ecis* (Deel 20, p. 1).

Technische Documentatie

- Algra, K., Bouter, L. M., Hol, A. & van Kreveld, J. (2018). Nederlandse gedragscode wetenschappelijke integriteit. Generic.
- Analytics, D. (2013). Open growth. stimulating demand for open data in the uk. Electronic Article. Verkregen van <https://www2.deloitte.com/content/dam/Deloitte/uk/Documents/deloitte-analytics/open-growth.pdf>
- Berners-Lee, T. (2006). Linked open data. Web Page. Verkregen van <https://www.w3.org/DesignIssues/LinkedData.html>
- Berners-Lee, T. (2012). 5 stars for open data. Web Page. Verkregen van <http://5stardata.info/>
- Digitale Overheid, M. v. B. (2020a). Data agenda overheid. Web Page. Verkregen van <https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/nieuwe-technologieen-data-en-ethiek/data-agenda-overheid/>
- Digitale Overheid, M. v. B. (2020b). Overheidsdata kwalitatief verbeteren en efficiënter benutten. Web Page. Verkregen van <https://www.digitaleoverheid.nl/overzicht-van-alle-onderwerpen/nieuwe-technologieen-data-en-ethiek/data-agenda-overheid/overheidsdata-kwalitatief-verbeteren-en-efficiënter-benutten/>
- Force11. (2020). Guiding principles for findable, accessible, interoperable and re-usable data. Web Page. Verkregen van <https://www.force11.org/group/fairgroup/fairprinciples>
- Gartner. (2020). Information technology glossary. Web Page. Verkregen van <https://www.gartner.com/en/information-technology/glossary>
- knaw. (2018). *Hergebruik van publieke data* (Report Nr. 978-90-6984-728-3). Koninklijke Nederlandse Akademie van Wetenschappen.
- Koninkrijksrelaties, D. M. v. B. Z. e. (2019). Tussenrapportage over de voortgang van de aanpak van de onvolkomenheden uitgesplitst per departement. Government Document. doi:26643
- KOOP, M. v. B. (2020). Koop overheid. Web Page. Verkregen van <https://www.koopoverheid.nl>
- KOOP, M. v. B. (2021a). Datacommunities van de nederlandse overheid. Web Page. Verkregen van <https://datacommunities.nl>
- KOOP, M. v. B. (2021b). Dataregister van de nederlandse overheid. Web Page. Verkregen van <data.overheid.nl>
- KOOP, M. v. B. (2021c). Wegwijzer naar informatie en diensten van alle nederlandse overheden. Web Page. Verkregen van <www.overheid.nl>
- Lyon, L., Ball, A. [Alex], Duke, M. & Day, M. (2012). Community capability model framework. Generic.
- NL, D. (2021). Dama nl. Web Page.
- NORA, M. v. B. (2020). Nora. Web Page. Verkregen van <https://www.noraonline.nl>
- Open Overheid, M. v. B. (2017). Open overheid en open data in het regeerakkoord. Web Page.
- Tauberer, J. & Lessig, L. (2020). The 8 principles of open government data. Web Page. Verkregen van <https://opengovdata.org>
- Yin Robert, K. (2017). Case study research and applications: Design and methods. Generic.

Acroniemen

- API** Application Programming Interface.
- BDM** Big Data Management.
- BI** Business Intelligence.
- BIR** Informatiebeveiliging Rijksdienst.
- BOLD** Big Open Linked Data.
- BPM** Business Process Management.
- BZK** Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties.
- CBR** Centraal Bureau Rijvaardigheidsbewijzen.
- CCMF** Community Capability Model Framework.
- CDDC** Community Driven Data Curation.
- CEO** Chief Executive Officer.
- CIO** Chief Information Officer.
- CM** Community Models.
- CPO** Chief Privacy Officer.
- DA** Data Analytics.
- DAMA-DMBOK2** Data Management Framework.
- DAO** Data Agenda Overheid.
- DGOO** Directoraat-generaal Overheids Organisatie van BZK.
- DGVBR** Directoraat-generaal Vastgoed en Bedrijfsvoering Rijks van BZK.
- DM** Data Management.
- DML** Data Management Lifecycle.
- DPO** Data Protection Officer.
- DSRM** Design Science Research Methode.
- DWH** Data WareHouse.
- DYA** DYnamic Architecture.
- EDM** Enterprise Data Model.
- EOSC** European Open Science Cloud.
- ERP** Enterprise Resource Planning.
- EU** Europese Unie.
- EUPL** European Union Public Licence.
- EZK** Nederlandse ministerie van Economische Zaken en Klimaat.
- GDPR** Europese General Data Protection Regulation.
- GEAF** Government Enterprise Architecture Framework.
- GGD** Gemeentelijke of Gemeenschappelijke Gezondheidsdienst.
- GGD-GHOR** Geneeskundige Hulpverleningsorganisatie in de Regio.
- ISO** International Organization for Standardization.
- IT** Informatietechnologie.
- ITG** IT-Governance.
- ITU-T** International Telecommunication Union - Telecommunications sector.
- JSON** JavaScript Object Notation.
- KNAW** Koninklijke Nederlandse Academie van Wetenschappen.
- KNMI** Koninklijk Nederlands Meteorologisch Instituut.
- KOOP** Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties.
- KPI** Kritieke Prestatie-Indicatoren.
- KT** Kepner Tregoe.
- LD** Linked Data.
- LOD** Linked Open Data.
- Logius** Logius.
- MD** Master Data.
- MDM** Master Data Management.
- MIT** Massachusetts Institute of Technology.
- MMDIS** Multidimensional Management and Development of Information Systems.
- MT** Managementteam.
- OCW** Nederlandse minister van Onderwijs, Cultuur en Wetenschap.
- ODP** Open Distributed Processing.
- OECD** Organisation for Economic Co-operation and Development.
- OGD** Open Government Data.
- OU** Open Universiteit.
- OWMS** Overheid.nl Web Metadata Standaard.
- PDOK** Publieke Dienstverlening Op de Kaart.
- PO** Product Owner.
- RDF** Resource Description Framework.

RM-ODP Reference Model for Open Distributed Processing.

SZW Nederlandse ministerie van Sociale Zaken en Werkgelegenheid.

TNO Toegepast-Natuurwetenschappelijk Onderzoek.

TOGAF The Open Group Architecture Framework.

Turtle Terse RDF Triple Language.

UBR Uitvoeringsorganisatie Bedrijfsvoering Rijk.

UDHR Universal Declaration of Human Rights.

URI Uniform Resource Identifier.

VNG Vereniging Nederlandse Gemeenten.

W3C World Wide Web Consortium.

Woordenlijst

Data Data zijn *gegevens*, die feiten en statistieken verzameld ter referentie of voor analyse.

Architectuur Met architectuur worden de fundamentele concepten/eigenschappen van een systeem in zijn omgeving, bestaande uit zijn deelsystemen, relaties en principes van zijn ontwerp en evolutie. Definitie: "Architectuur: fundamental concepts or properties of a system in its environment, embodied in its elements, relationships, and in the principles of its design and evolution" ([Lankhorst, 2009]) .

AS-IS Definitie: 'AS-IS' beschrijft een proces of situatie zoals deze nu functioneert .

AVG Uitvoeringswet Algemene Verordening Gegevensbescherming (AVG)
<https://wetten.overheid.nl/BWBR0040940/2021-07-01>.

BD Men spreekt van "**Big Data**" (BD) wanneer men werkt met een of meer datasets die te groot zijn om met reguliere databasemanagementsystemen onderhouden te worden.

BPM&IT Master opleiding van de Open Universiteit (OU) op het kruisingsvlak van Business Process Management (BPM) en Informatietechnologie (IT).

Capaciteitsfactoren Capaciteitsfactoren ofwel "capability factors" kan worden beschreven in termen van: bekwaamheid, geschiktheid, kundig en in staat zijn tot.
<https://nl.bab.la/woordenboek/engels-nederlands/capabilities>.

CBS Het Centraal Bureau voor de Statistiek (CBS) verzamelt, bewerkt en publiceert statistische gegevens over de Nederlandse samenleving.
<https://www.cbs.nl>.

DAMA Staat voor Data Management en is een non-profitorganisatie voor globale Data Management communicatie.
<https://dama.org>.

DAMA-DMBOK Data Management Body Of Knowledge [Henderson, 2017].

DAMA-NL De Nederlandse afdeling van DAMA see.
<https://dama-nl.org>.

data.overheid.nl data.overheid.nl is het Dataregister van de Nederlandse Overheid [KOOP, 2021b].
<https://data.overheid.nl>.

datacommunities.nl datacommunities.nl is het platform ter ondersteuning van data communities van de Nederlandse Overheid [KOOP, 2021a].
<https://datacommunities.nl>, vanaf 2023: <https://data.overheid.nl/datacommunities>.

DBOU De online Digitale Bibliotheekservice van de Open Universiteit.
<https://bibliotheek.ou.nl>.

DCAT DCAT is een RDF vocabulaire dat is ontworpen om de interoperabiliteit tussen op het web gepubliceerde gegevenscatalogi te vergemakkelijken.
<https://www.w3.org/TR/vocab-dcat-2/>.

DG Er zijn verschillende definities van Data Governance (DG): Definitie: "A companywide framework for assigning decision-related rights and duties in order to be able to adequately handle data as a company asset. " ([Otto, 2011a; Otto, 2011b]) Definitie: "Data Governance (DG) is the specification of decision rights and an accountability framework to ensure the appropriate behavior in the valuation, creation, consumption and control of data and analytics. " ([Gartner, 2020]) .

- DO** directie Digitale Overheid (DO) van DG00, zorgt ervoor dat burgers en bedrijven op een veilige, transparante en betrouwbare manier digitaal hun zaken met de overheid kunnen doen.
<https://www.digitaleoverheid.nl>.
- DUO** Dienst Uitvoering Onderwijs (DUO) voert als baten-lastendienst in opdracht van de OCW onderwijswet-ten en -regelingen uit. Ook voert DUO in opdracht van het ministerie van SZW de Wet inburgering uit.
<https://duo.nl/organisatie/organisatie>.
- EA** Enterprise Architectuur (EA) is een samenhangend geheel van principes, methoden en modellen die worden gebruikt bij het ontwerpen en realiseren van de organisatiestructuur, bedrijfsprocessen, informatie-systemen en infrastructuur van een onderneming. Definitie: "Enterprise Architectuur (EA): a coherent whole of principles, methods, and models that are used in the design and realisation of an enterprise's organisational structure, business processes, information systems, and infrastructure." ([Lankhorst, 2009]) Definitie: "Enterprise Architectuur (EA) is a description of an enterprise from an integrated business and IT perspective" ([Kotusev, 2017]).
- EAM** Enterprise Architecture Management (EAM) is een conceptueel model voor gedecentraliseerde onder-nemingen. Definitie: "Enterprise Architecture Management (EAM) is a management practice embracing all the management processes related to EA aiming to improve business and IT alignment" ([Kotusev, 2017]).
- Enterprise** Enterprise is een verzameling organisaties met een gemeenschappelijke set doelen. Definitie: "Enterprise: any collection of organisations that has a common set of goals and/or a single bottom line." ([Lankhorst, 2009]).
- EUR** Erasmus Universiteit Rotterdam (EUR)
<https://www.eur.nl>.
- FADAM** Het FADAM-framework wordt voorgesteld als een eerste aanzet van een aanpak voor FAIR Data Management. Definitie: "The FAIR Data Management (FADAM) framework is proposed, as an early stage of an approach to FAIR Data Management" ([dFigueiredo, 2019]).
- FAIR** FAIR, of FAIR Principles, of FAIR Data Principles, zijn een set principes [Wilkinson, 2016]: findable, accessible, interoperable, reusable.
<https://www.go-fair.org/fair-principles>.
- FAIR Data** FAIR Data.
- FAIR Data Principles** FAIR Data Principles.
- GO FAIR** GO FAIR Initiative.
<https://www.go-fair.org>.
- GS** Google Scholar biedt een eenvoudige manier om breed te zoeken naar wetenschappelijke literatuur.
<https://scholar.google.com>.
- Kadaster** Het Kadaster registreert gegevens van registergoederen: onroerende zaken (grond, gebouwen en kabelnetten), schepen en luchtvaartuigen.
<https://www.kadaster.nl>.
- Leiden** Universiteit Leiden (Leiden)
<https://www.universiteitleiden.nl>.
- Metadata** Metadata beschrijven de karakteristieken van de data waarop het betrekking heeft.
- NORA** Nederlandse Overheid Referentie Architectuur (NORA)
<https://www.noraonline.nl>.
- Open data** Er zijn acht uitgangspunten gedefinieerd voor het beschikbaar stellen van **open data** van de overheid: (1) Afwegingskader voor openbaarheid. (2) Open data is gratis. (3) Open data is vrij van rechten

van derden. (4) Open data is zonder registratie toegankelijk. (5) Open data is computer verwerkbaar. (6) Open data is voorzien van metadata. (7) Open data is zo volledig en onbewerkt mogelijk. (8) Open data is vindbaar. Definitie:

Open data is information or content made freely available to use and redistribute, subject only to the requirement to attribute it to the source. The term also may be used more casually to describe any data that is shared outside the organization and beyond its original intended use, for example, with business partners, customers or industry associations. Formally, data designated as "open" is subject to several conditions and licensing that can be found at opendefinition.org. ([Gartner, 2020])

<https://data.overheid.nl/ondersteuning/open-data/wat-is-open-data>.

OSF Open State Foundation (OSF), is een onafhankelijke stichting die zich inzet voor een Digitale transparantie en een open overheid.

<https://openstate.eu>.

PA Probleemanalyse (SB), is een van de denkprocessen van Kepner en Tregoe [Kepner, 1997].

Deze bestaat uit de volgende stappen: (i) Beschrijf Probleem, (ii) Formuleer Mogelijke Oorzaken, (iii) Evalueer Mogelijke Oorzaken, (iv) Bewijs de Juiste Oorzaak. Problem Analysis (PA) of Kepner and Tregoe [Kepner, 1997, Chap. Two]

<https://www.kepner-tregoe.com/about-kt/thinking-as-a-competitive-advantage/>.

PLOOI Platform Open Overheidsinformatie (PLOOI)

<https://www.koopoverheid.nl/voor-overheden/gemeenten-provincies-en-waterschappen/plooi>.

publieke overheidsdata Publieke overheidsdata, bestaat uit open data en semi-open data..

RIG Rijks ICT Gilde (RIG) van het ministerie van BZK, Rijks tech consultants voor het Nederland van morgen.

<https://www.rijksictgilde.nl>.

SB Situatiebeoordeling (SB), is een van de denkprocessen van Kepner en Tregoe [Kepner, 1997].

Deze bestaat uit de volgende stappen: (i) Identificeer Aandachtspunten, (ii) Stel Prioriteiten, (iii) Bepaal de Volgende Stappen, (iv) Plan Betrokkenheid. Situation Appraisal (AS) of Kepner and Tregoe [Kepner, 1997, Chap. Seven]

<https://www.kepner-tregoe.com/about-kt/thinking-as-a-competitive-advantage/>.

Semi-open data De volgende 8 uitgangspunten definiëren het kader voor dit onderzoek voor het beschikbaar stellen van **semi-open data** van de overheid. (1) Afwegingskader voor openbaarheid. (2) Semi-open data hoeft niet gratis te zijn. (3) Semi-open data hoeft niet vrij van rechten van derden te zijn. (4) Semi-open data kan alleen via registratie toegankelijk zijn. (5) Semi-open data kan computer verwerkbaar zijn. (6) Semi-open data kan zijn voorzien van metadata. (7) Semi-open data is zo volledig en onbewerkt mogelijk. (8) Semi-open data is niet altijd vindbaar..

SFA De drie waarderingsfactoren zijn: (i) Suitability (geschiktheid), (ii) Feasibility (haalbaarheid), (iii) Acceptability (aanvaardbaarheid)..

SMART Definitie: (i) Specifiek - Is de doelstelling eenduidig? (ii) Meetbaar - Onder welke (meetbare/observeerbare) voorwaarden of vorm is het doel bereikt? (iii) Acceptabel - Zijn deze doelen acceptabel voor de doelgroep en/of het management? (iv) Realistisch - Is het doel haalbaar? (v) Tijdsgebonden - Wanneer (in de tijd) moet het doel bereikt zijn?

<https://nl.wikipedia.org/wiki/SMART-principe>.

TO-BE Definitie: Een 'TO-BE' proces is de beschrijving van een proces. De beschrijving geeft aan hoe het proces gaat functioneren na doorvoering van een aantal voorgenomen verbeteringen.

TUD Technische Universiteit Delft (TUD)

<https://www.tudelft.nl>.

UvA Universiteit van Amsterdam (UvA)

<https://www.uva.nl/>.

Who Wet hergebruik van overheidsinformatie (Who)

<https://wetten.overheid.nl/BWBR0036795/2021-07-01>.

Wob Wet openbaarheid van bestuur (Wob)

<https://wetten.overheid.nl/BWBR0005252/2018-07-28>.

Woo Wet open overheid (Woo)

Vervangt de Wob per 1 mei 2022.

<https://wetten.overheid.nl/BWBR0045754/2022-08-01>.

www.digitaleoverheid.nl Digitale Overheid slaat een brug tussen beleid en professionals die werken aan digitalisering van de overheid.

<https://www.digitaleoverheid.nl>.

www.overheid.nl www.overheid.nl is de wegwijzer naar informatie en diensten van alle overheden [KOOP, 2021b].

<https://www.overheid.nl>.

Lijst van figuren

1.1	Gebiedsverkenning	2
1.2	Publieke overheidsdata	3
1.3	Onderzoekstappen	6
2.1	Vereenvoudigde conceptueel model van Enterprise Architecture Management (EAM) voor ge- decentraliseerde ondernemingen	12
2.2	Volwassenheidsniveau van open data governance in Nederland in 2014	13
2.3	Conceptueel model	16
3.1	Onderzoeksmodel	20
4.1	Organigram casus-organisatie	25
4.2	Analyse kernbegrippen	35
A.1	5 sterren stappen	52
A.2	Conceptueel model van EAM voor gedecentraliseerde ondernemingen	57
B.1	Interview DO	75

Lijst van tabellen

1.1	Onderzoeksvragen voor het literatuuronderzoek	5
1.2	Onderzoeksvragen voor het empirisch onderzoek.	6
2.1	Uitvoering literatuuronderzoek	9
2.2	Beantwoording onderzoeksvragen voor het empirisch onderzoek	16
2.3	Kwalitatieve informatiebronnen voor de verschillende dimensies	17
3.1	Plan van aanpak	21
4.1	Respondenten en onderzoeksvragen	26
4.2	Samenvatting documentanalyse	34
A.1	Vijf elementen van volwassenheid van Open Government Data (OGD)	59
A.2	OGD volwassenheidsmodel	60
A.3	Master Data Management (MDM) vijf beslisdomeinen	61
A.4	FAIR leidende principes	62
A.5	FAIR Data Management (FADAM) FAIR raamwerk	63
A.6	<i>Best Practices</i> Community Capability Model Framework (CCMF)	64
A.7	<i>Best Practices</i> Community Driven Data Curation (CDDC)	65
A.8	Beantwoording literatuurvragen	66
A.8	Beantwoording literatuurvragen	67
A.9	Kernbegrippen, dimensies en indicatoren	68
A.9	Kernbegrippen, dimensies en indicatoren	69
A.9	Kernbegrippen, dimensies en indicatoren	70
A.9	Kernbegrippen, dimensies en indicatoren	71
C.1	Template transcriptie interview	77
C.1	Template transcriptie interview	78
C.1	Template transcriptie interview	79
C.1	Template transcriptie interview	80
C.2	Template analyse document	81
D.1	Analyse kernbegrippen	92
D.2	Analyse document Wet openbaarheid van bestuur	93
D.3	Analyse document Wet open overheid	94
D.4	Analyse document Wet hergebruik van overheidsinformatie	95
D.5	Analyse document Data Catalog Vocabulary	96

A

Verdieping

In deze bijlage is de verdieping van enkele onderwerpen van dit onderzoek verzameld. Deze hebben betrekking op de volgende onderwerpen:

- Gebiedsverkenning, paragraaf A.1
- Theoretische kader, paragraaf A.2

A.1. Gebiedsverkenning

In de gebiedsverkenning van hoofdstuk '1-Introductie', paragraaf '1.2-Gebiedsverkenning' zijn aan aantal onderwerpen de revue gepasseerd. Van enkele van deze onderwerpen is een toelichting of verdieping gegeven.

- Publieke overheidsdata paragraaf 1.2.3
- Hergebruikers paragraaf A.1.2

A.1.1. Publieke overheidsdata

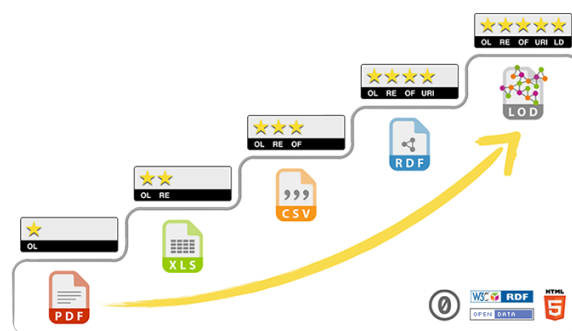
De publieke overheidsdata zijn onder te verdelen in open data, waarvan er een duidelijke definitie is, en semi-open data waarvan er geen duidelijke definitie is. In dit onderzoek maken we de onderstaande onderschijft in open data en semi-open data, daarmee volgen we de aanbevelingen van Bargh e.a. [Bargh, 2016].

Open data

Wat is open data precies? Deze acht uitgangspunten definiëren het kader voor het beschikbaar stellen van open data van de overheid. (1) Afwegingskader voor openbaarheid. (2) Open data is gratis. (3) Open data is vrij van rechten van derden. (4) Open data is zonder registratie toegankelijk. (5) Open data is computer verwerkbaar. (6) Open data is voorzien van metadata. (7) Open data is zo volledig en onbewerkt mogelijk. (8) Open data is vindbaar. Zie paragraaf 'A.1.1-Open data' voor een gedetailleerde beschrijving.

Standaarden rond open data worden binnen de overheid geborgd in de Nederlandse Overheid Referentie Architectuur (NORA), hiervoor wordt Linked Data (LD) als raamwerk gebruikt [NORA, 2020]. Het volwassenheidsniveaus van open data, weergegeven met het aantal sterren, zijn schematisch weergegeven in figuur 'A.1-5 sterren stappen'.

Hergebruikers op basis van de rol die een gebruiker van open data speelt in de open data waardeketen, zijn te verdelen in, Verzamelaars, Enablers, Ontwikkelaars, Verrijkers en Eindgebruikers [Analytics, 2013; Donker,



<https://5stardata.info/en/>

Figuur A.1: 5 sterren stappen

2019], zie bijlage 'A-Verdieping' paragraaf A.1.2 voor een verdere beschrijving van deze hergebruikers. Binnen dit onderzoek houden we de onderstaande definitie aan van open data, die onderdeel is van de publieke overheidsdata.

Deze acht uitgangspunten definiëren het kader voor het proactief beschikbaar stellen van open data van de overheid¹.

1. Afwegingskader voor openbaarheid

Data van een overheidsorganisatie wordt proactief beschikbaar gesteld als open data indien geen uitzonderingsgronden of beperkingsgronden van de Wet openbaarheid van bestuur of bijzondere openbaarmakingsregelingen van toepassing zijn. Bij het beschikbaar stellen van open data dient een afweging gemaakt te worden of hergebruik risico's oplevert voor fundamentele waarden en privacy.

2. Open data is gratis

Overheidsorganisaties brengen geen kosten in rekening voor het beschikbaar stellen van open data.

3. Open data is vrij van rechten van derden

Open data wordt gevrijwaard van rechten van derden door vrijgave met een CC-0 verklaring. Indien de vermelding van de bron van de data van belang is voor een overheidsorganisatie, kan open data worden vrijgegeven met een CC-BY verklaring.

4. Open data is zonder registratie toegankelijk

Open data is toegankelijk zonder dat er enige vorm van registratie van gegevens van de potentiële gebruiker plaatsvindt.

5. Open data is computer verwerkbaar

Open data wordt aangeboden op een manier die verdere verwerking met een computer mogelijk maakt. Bij voorkeur wordt Open data beschikbaar gesteld in een open standaard formaat.

6. Open data is voorzien van metadata

Open data is voorzien van metadata conform de DCAT-AP standaard.

7. Open data is zo volledig en onbewerkt mogelijk

Data wordt in een voor hergebruik geschikte vorm door een overheidsorganisatie beschikbaar gesteld. Open data is qua kwaliteit en actualiteit zo gelijk mogelijk aan de binnen de publieke organisatie gebruikte data en wordt zoveel mogelijk 'as-is' beschikbaar gesteld. Data kan bewerkt worden om te voldoen aan het afwegingskader voor openbaarheid zolang deze bewerking niet in strijd is met de mededingingswet.

8. Open data is vindbaar

Overheidsorganisaties maken hun open data vindbaar door op <https://data.overheid.nl> een verwijzing naar de door hen beschikbaar gestelde open data te geven.

Semi-open data

Wat is semi-open data precies?

De volgende 8 uitgangspunten definiëren het kader voor dit onderzoek voor het beschikbaar stellen van semi-open data van de overheid. (1) Afwegingskader voor openbaarheid. (2) Semi-open data hoeft niet gratis te zijn. (3) Semi-open data hoeft niet vrij van rechten van derden te zijn. (4) Semi-open data kan via registratie toegankelijk zijn. (5) Semi-open data kan computer verwerkbaar zijn. (6) Semi-open data kan zijn voorzien van metadata. (7) Semi-open data is zo volledig en onbewerkt mogelijk. (8) Semi-open data is niet altijd vindbaar. Er zijn geen standaarden voor semi-open data opgesteld [Bargh, 2016].

Deze acht uitgangspunten definiëren het kader voor het proactief beschikbaar stellen van open data van de overheid [KOOOP, 2021b].

1. Afwegingskader voor openbaarheid

Het afwegingskader voor openbaarheid van semi-open data is vergelijkbaar met die van open data en voor alle publieke overheidsdata. Hieraan kan zijn aangegeven onder welke voorwaarden deze openbaarheid geldt.

¹ <https://data.overheid.nl/ondersteuning/open-data/wat-is-open-data>

2. **Semi-open data hoeft niet gratis te zijn**
Overheidsorganisaties kunnen kosten in rekening voor het beschikbaar stellen van semi-open data, als dat nodig is om aan specifieke (publieke) publieke overheidsdata vragen te kunnen voldoen.
3. **Semi-open data hoeft niet vrij van rechten van derden te zijn**
Als het gaat om specifieke (publieke) publieke overheidsdata dan kunnen daar nog rechten van derden op liggen, zie afwegingskader.
4. **Semi-open data kan via registratie toegankelijk**
Als het gaat om specifieke publieke overheidsdata kan nodig zijn eerst te moeten registreren.
5. **Semi-open data kan computer verwerkbaar zijn**
Niet alle publieke overheidsdata kan direct worden verwerkt door computers, maar kan wel openbaar zijn.
6. **Semi-open data kan zijn voorzien van metadata**
Gezien Semi-open data niet altijd digitaal kan deze niet altijd zijn voorzien van metadata.
7. **Semi-open data is zo volledig en onbewerkt mogelijk**
Dit geldt voor open data maar is niet anders als voor semi-open data en geldt dus ook voor publieke overheidsdata.
8. **Open data is niet altijd vindbaar**
Het doel moet zijn wel vindbaar te zijn, maar voor specifieke (publieke) publieke overheidsdata kunnen er andere afspraken voor zijn gemaakt.

A.1.2. Hergebruikers

Het is niet mogelijk om alle hergebruikers van publieke overheidsdata bij dit onderzoek te betrekken, zoals, Verzamelaars, Enablers, Ontwikkelaars, Verrijkers en Eindgebruikers [Donker, 2019, p. 26-27]. De Product Owner (PO) van Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) heeft aangegeven graag te willen meewerken aan dit onderzoek, mede omdat zij de taakstelling hebben het hergebruik van open data te verbeteren. Daarnaast onderhoudt KOOP het contact met verschillende groepen hergebruikers van open data.

Hergebruikers van publieke overheidsdata ofwel open data en semi-open data spelen een belangrijke rol binnen dit onderzoek. Binnen dit onderzoek worden de definities gebruik van Analytics [Analytics, 2013] en Donker e.a. [Donker, 2019]. De rol die een gebruiker van open data speelt in de open data waardeketen zijn onder te verdelen in: Verzamelaars, Enablers, Ontwikkelaars, Verrijkers en Eindgebruikers.

Verzamelaars

Verzamelaars zijn organisaties die open data verzamelen en aggregeren, al dan niet in combinatie met semi-open data. Dit verzamelen en/of aggregeren van data vindt vaak plaats op sectoren of op geografisch niveau. Door het verzamelen en/of aggregeren van data uit verschillende bronnen kunnen de data worden gebruikt voor bijvoorbeeld het efficiënt aanbieden van data, of voor het uitvoeren van analyses (door verzamelaars zelf of door andere gebruikers) [Donker, 2019, p. 26].

Enablers

Enablers zijn organisaties die een platform en technologieën aanbieden zodat derden gebruik kunnen maken van open data. De enablers maken zelf geen gebruik van de data, maar bieden kosteneffectieve en efficiënte oplossingen aan, zowel aan aanbieders als aan (her)gebruikers van open data. Enablers verwerken de ruwe informatie in een dienstenplatform waar verzamelaars van meerdere domeinen worden samengebracht. Naast het beschikbaar stellen van open data, kan een enabler aanvullende diensten beschikbaar stellen zoals communicatie, advies, maatwerkdiensten en trainingen. Enablers kunnen ook als makelaars tussen data-aanbieders en gebruikers fungeren, bijvoorbeeld door te lobbyen voor het publiceren van data die nog niet als open data beschikbaar is [Donker, 2019, p. 26].

Ontwikkelaars

Ontwikkelaars zijn organisaties en softwareondernemers die tools en toepassingen ontwikkelen om de data nog beter te gebruiken. Ze ontwerpen, bouwen en verkopen bijvoorbeeld web- en slimme telefoontoepassingen (vaak gebaseerd op dynamische data). De afnemers van de diensten van de ontwikkelaars kunnen consumenten zijn maar ook overheden omdat open data aantrekkelijker en informatiever kan worden aangeboden. De ontwikkelaars onderscheiden zich van de enablers doordat zij wel direct gebruik maken van open data [Donker, 2019, p. 27].

Verrijkers

Verrijkers verwerken de ruwe informatie tot nieuwe producten, al dan niet met behulp van ontwikkelaars. Dit zijn organisaties die open data veelal gebruiken om hun bestaande diensten/producten te verbeteren om nieuwe of betere inzichten kunnen leveren. Dergelijke producten/diensten zouden zonder open data ook kunnen bestaan, maar kunnen nu efficiënter en effectiever worden ingezet, bijvoorbeeld door verzekeraars of ingenieursbureaus [Donker, 2019, p. 27].

Eindgebruikers

Eindgebruikers zijn de gebruikers van de producten en diensten die op basis van open data zijn ontwikkeld. Dit kunnen zowel consumenten zijn als het bedrijfsleven en de overheid. Eindgebruikers kunnen ook weer aan het begin van een nieuwe waardeketen staan [Donker, 2019, p. 27].

A.1.3. Motivatie / relevantie

Vanuit het directie Digitale Overheid (DO) is aangegeven dat er al onderzoek is gedaan op dit gebied door de onderstaande wetenschappers:

1. Technische Universiteit Delft (TUD)
 - (a) Anneke Zuiderwijk [Zuiderwijk, 2014] ²
 - (b) Bastiaan van Loenen ³
 - (c) Frederika Welle Donker (gepromoveerd op open data) [Welle Donker, 2017] ⁴
2. Erasmus Universiteit Rotterdam (EUR)
 - (a) Pearl Dykstra (KNAW-rapport) ⁵
3. Universiteit Leiden (Leiden)
 - (a) Barend Mons [Mons, 2011]

Regeerakkoord 2017–2021

Het kabinet heeft in het regeerakkoord van 2017–2021 plannen opgenomen over de open overheid en open data, onder 'Vertrouwen in de toekomst' [Open Overheid, 2017]:

- 'Het kabinet hecht eraan dat de overheid transparant en open is. Er is een initiatiefvoorstel open overheid aanhangig. Er wordt onderzocht hoe de verruiming van de openheid gestalte kan krijgen zonder hoge kosten voor de organisatie en uitvoering. Het kabinet treedt daartoe in overleg met de initiatiefnemers.'
- 'De overheid beschikt over veel algemene, openbare informatie. Deze data worden goed vindbaar en toegankelijk gemaakt, in de vorm van open data.'

Hierbij moet worden gedacht aan publieke gegevens waarvoor een vergoeding of authenticatie noodzakelijk is. Het DO wil graag bij dit onderzoek betrokken worden en verleent hierbij alle medewerking.

²Anneke Zuiderwijk: <https://scholar.google.nl/citations?user=qmAD2KcAAAAJ&hl=nl&oi=ao>

³Bastiaan van Loenen: <https://scholar.google.nl/citations?user=50f8eK0AAAAJ&hl=nl&oi=ao>

⁴Frederika Welle Donker: <https://scholar.google.nl/citations?user=s580f8QAAAAJ&hl=nl&oi=ao>

⁵Pearl Dykstra: https://scholar.google.nl/citations?user=LB5T-_0AAAAJ&hl=nl&oi=ao

A.2. Theoretische kader

In het Theoretische kader van hoofdstuk '2-Theoretisch kader' is de literatuur onderzocht op het gebied van Enterprise Architectuur (EA) en Data Governance (DG) om daarmee de sturing op Big Data (BD) te kunnen verbeteren. Van enkele van deze onderwerpen is een toelichting of verdieping gegeven.

- Enterprise Architectuur paragraaf 2.3.2
- Data Governance paragraaf 2.3.3
- FAIR Data Principles en FAIR Data Management paragraaf 2.3.4
- Community Models paragraaf 2.3.5

In de woordenlijst zijn verschillende definities opgenomen waaronder die van Gartner [Gartner, 2020].

A.2.1. Enterprise Architectuur Principles

In de woordenlijst zijn definities van EA van Kotusev [Kotusev, 2017] en Lankhorst [Lankhorst, 2009].

Enterprise Architecture Management

Hieronder zijn de Enterprise Architecture Management (EAM) principes van Kotusev [Kotusev, 2017] beschreven. In figuur A.2 is het conceptueel model van EAM weergegeven [Kotusev, 2017].

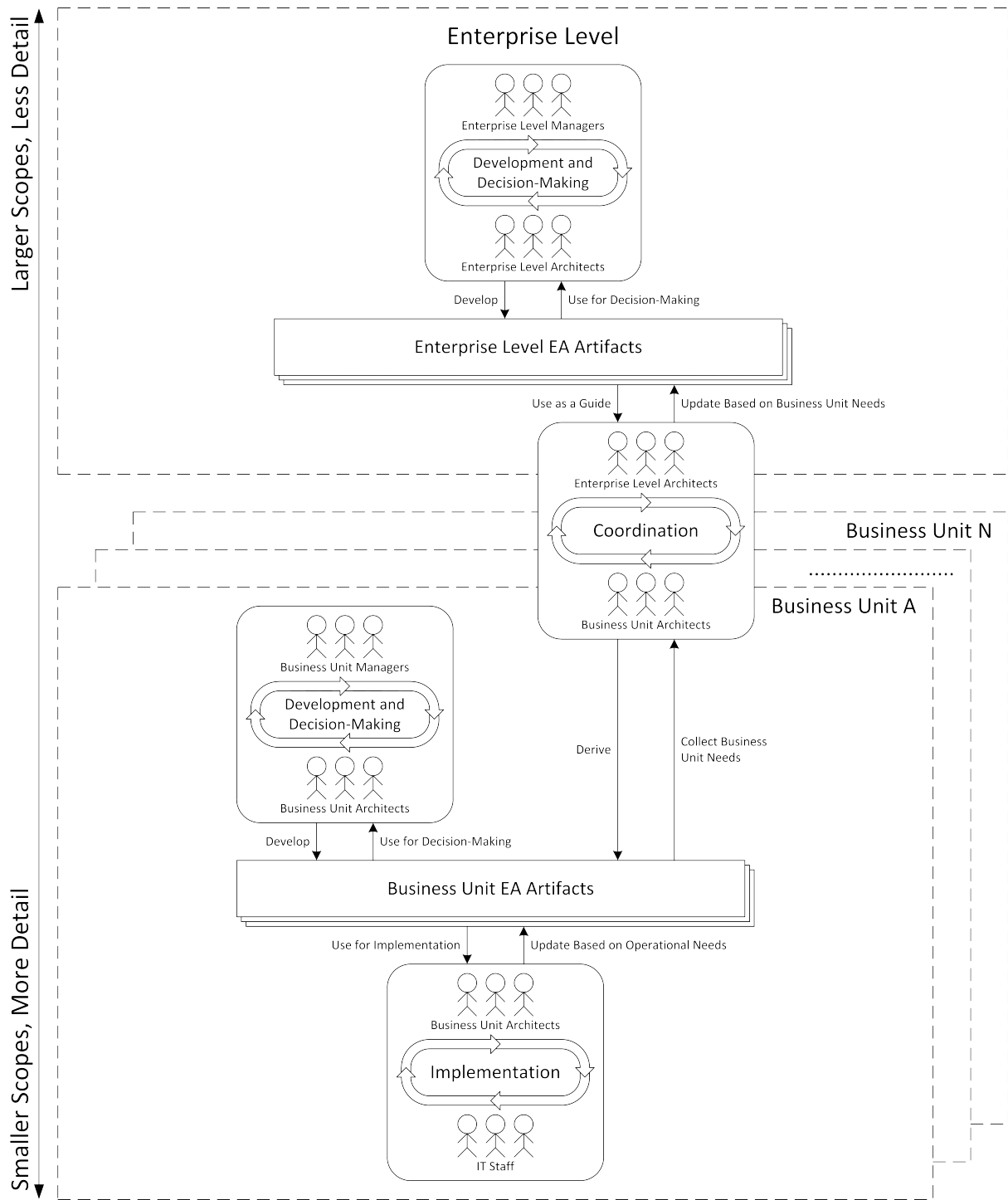
Elements of the conceptual model of EAM

- Decision-making
 - ◇ Managers and architects use corresponding EA artifacts (enterprise level or business unit) for making strategic, IT investments or other important decisions.
- Coordination
 - ◇ Enterprise level and business unit architects collaborate to balance local and global needs and to achieve a shared architectural vision.

The influence of "soft" human factors on EAM

- ALL EAM processes
 - ◇ Overall organizational acceptance of EAM.
Overall organizational acceptance of EAM, determined by enforcement, rewards, support, management commitment, social legitimacy, performance expectancy, training, business buy-in, organizational politics and other factors, significantly influences the success of EAM.
 - ◇ General features of the organizational culture.
General features of the organizational culture (including leadership support, awareness among stakeholders, common understanding of EA, orientation to collaboration, control and responsibility, orientation to work, internal vs external focus and change vs stability) can have a considerable and complex impact on EAM.
- Development and decision-making processes in all enterprises
 - ◇ Cultural differences between architects and managers.
The cultural differences in seven dimensions of culture (basis of truth, nature of time, orientation to work, orientation to change, control and responsibility, orientation and focus and orientation to IT) can cause communication breakdowns between architects and managers and influence on EAM.
- Coordination processes in decentralized enterprises
 - ◇ Cultural differences between different architects.
The cultural differences in seven dimensions of culture (basis of truth, nature of time, orientation to work, orientation to change, control and responsibility, orientation and focus and orientation to IT) can cause communication breakdowns between architects from different architecture teams and influence on EAM.

Het volledige conceptueel model van EAM zoals Kotusev [Kotusev, 2017] heeft beschreven, is weergegeven in figuur 'A.2-Conceptueel model van EAM voor gedecentraliseerde ondernemingen'.



Figuur A.2: Conceptueel model van EAM voor gedecentraliseerde ondernemingen [Kotusev, 2017, p. 16]

Government Enterprise Architecture Framework

Hieronder zijn de zes Government Enterprise Architecture Framework (GEAF) principes van Lnenicka en Kmarkova [Lnenicka, 2019] beschreven.

Functional suitability

- Functional completeness
 - ◇ The execution of requirements for the functional complete Architecture provides all the functions according to the access rights defined for the data user role.
- Functional correctness
 - ◇ The execution of requirements for the functional correctness Architecture provides all the functions according to the requirements of the open (big) data lifecycle.

Reliability

- Availability
 - ◇ System availability Architecture of the system requires the availability of 99%, excluding scheduled maintenance times. The highest priority has the data infrastructure.
- Fault tolerance
 - ◇ Data storing and processing faults Distributed architecture that uses data replication techniques is used to provide fault tolerance.

Performance efficiency

- Time behaviour
 - ◇ The number of requests (throughput raes) Architecture uses distributed data processing and scalable cloud resources to solve the growing demands in time.

Usability

- User error protection
 - ◇ Errors (problems) and their solutions The ratio of the number of user-reported and subsequently solved problems is required to be more than 90%.

Security

- Confidentiality
 - ◇ Authorization Authorization processes are supported by the architecture to gain access to data. These are required through all the lifecycle phases except data publication.
 - ◇ Anonymization Data anonymization processes are supported by the architecture in the data management and preparation phase.
- Integrity
 - ◇ Data integrity Architecture supports mechanisms to protect against unauthorized access and assure the accuracy and consistency of data.

Compatibility

- Interoperability
 - ◇ Support, management and control Architecture supports this at the technical (protocols), semantic (data description standards), and organizational level.

A.2.2. Data Governance Principles

In de woordenlijst is definitie van DG van Otto [Otto, 2011a; Otto, 2011b].

Open Government Data

In tabel A.1 zijn de vijf elementen van volwassenheid van Welle Donker en van Loenen [Welle Donker, 2017] weergegeven.

Vijf elementen van volwassenheid van Open Government Data (OGD)	
Vision	to provide a common goal, to avoid a fragmented approach and to stimulate cooperation between stakeholders.
Leadership & Control	open data need to have a problem owner who will stimulate and coordinate open data activities. Awareness creation and capacity-building may lead to political support for open data, which is an important success factor, as is work floor support.
Financing Open Data	should extend beyond the initial stages and become embedded in the organisation's budget for data management and infrastructures.
Self-Organising ability	the way in which supply matches demand. In the initial stages, it will be mostly data providers requiring answers to specific questions and proactively promote open data. In later stages, matching supply and demand will increasingly become a part of the organisation's culture.
Communication channels	with whom, how, and what is communicated. In the initial stages, this will be mostly internal communication and in later stages, also external communication.

Tabel A.1: OGD [Welle Donker, 2017]

Open Government Data volwassenheidsmodel

In tabel A.2 is het volwassenheidsmodel van Welle Donker en van Loenen [Welle Donker, 2017] weergegeven.

OGD volwassenheidsmodel					
Organisatorische ontwikkelingsfasen	Fase I (standalone)	Fase II	Fase III	Fase IV	Fase V (netwerk)
Visie	Visie binnen eigen afdeling maar niet voor hele organisatie	Gedeelde visie binnen organisatie	Gedeelde visie rijksoverheid	Gedeelde visie gehele overheid	Gedeelde visie met politiek draagvlak
Leiderschap	Leiderschap op afdelingsniveau, regie op afdelingsniveau	Nationaal leiderschap moet zich nog uitkristalliseren, regie op organisatieniveau	Leiderschap met organisatieoverstijgende regie binnen overheid	Geaccepteerde regie binnen overheid & internationale beïnvloeding	IV+ regie mede bepalend op internationaal niveau
Financiering open data	Geen financiering	Eenmalige financiering/ subsidie voor pilot	Jaarlijkse financiering	Meer jaren programma budget	Gegarandeerde financiering
Zelf-ontwikkelen vermogen	Geen stimulering, wel vereist	Eenmalige stimulering zonder opvolging	Periodieke stimulering maar eventgedreven (fragmentarisch)	structureel vraag koppelen aan aanbod (platform, forum)	open data geïnstitutionaliseerd (geen stimulering nodig)
Communicatie S2U	Geen communicatie	Data passief aangeboden met alleen algemene contactgegevens	Aanbodgericht 1-weg communicatie naar gebruikers: "volg de aanbieder"	III + Vraaggerichte communicatie: Formeel tussen aanbieders en gebruikers	IV + "Volg de gebruiker": Vraaggerichte 2-wegs communicatie: Formeel aanbieders-gebruikers op alle niveaus:
Communicatie G2G	Communicatie 1-weg binnen eigen organisatie	Communicatie 2-weg binnen eigen organisatie	Communicatie tussen overheden op dezelfde overheidslaag (informeel)	Communicatie tussen overheden op dezelfde overheidslaag (formeel)	Communicatie tussen alle overheden (informeel + formeel)

Tabel A.2: OGD volwassenheidsmodel [Welle Donker, 2017]

Master Data Management

In tabel A.3 zijn de vijf beslisdomeinen van Khatri en Brown [Khatri, 2010] weergegeven.

Vijf principes van Master Data Management (MDM)	
Data Governance Domains	Domain Decisions
Data principles <ul style="list-style-type: none"> Clarifying the role of data as an asset 	<ul style="list-style-type: none"> What are the uses of data for the business? What are the mechanisms for communicating business uses of data on an ongoing basis? What are the desirable behaviors for employing data as assets? Who are opportunities for sharing and reuse of data identified? How does the regulatory environment influence the business uses of data?
Data Quality <ul style="list-style-type: none"> Establishing the requirements of intended use of data 	<ul style="list-style-type: none"> What are the standards for data quality with respect to accuracy, timeliness, completeness and credibility? What is the program for establishing and communicating data quality? How will data quality as well as the associated program be evaluated?
Metadata <ul style="list-style-type: none"> Establishing the semantics or "content" of data so that it is interpretable by the users 	<ul style="list-style-type: none"> What is the program for documenting the semantics of data? How will data be consistently defined and modeled so that it is interpretable? What is the plan to keep different types of metadata up-to-date?
Data Access <ul style="list-style-type: none"> Specifying access requirements of data 	<ul style="list-style-type: none"> What is the business value of data? How will risk assessment be conducted on an ongoing basis? How will assessment results be integrated with the overall compliance monitoring efforts? What are data access standards and procedures? What is the program for periodic monitoring and audit for compliance? How is security awareness and education disseminated? What is the program for backup and recovery?
Data Lifecycle <ul style="list-style-type: none"> Determining the definition, production, retention and retirement of data 	<ul style="list-style-type: none"> How is data inventoried? What is the program for data definition, production, retention, and retirement for different types of data? How do the compliance issues related to legislation affect data retention and archiving?

Tabel A.3: MDM [Khatri, 2010]

A.2.3. FAIR Data Management en FAIR Data Principles

GO FAIR is een initiatief dat ondersteuning en coördinatie biedt aan individuen, instellingen en organisaties die zich inzetten voor het FAIR maken van onderzoeksgegevens, volgens de aanbevelingen van de High Level Expert Group van de European Open Science Cloud (EOSC).

FAIR Data Principles

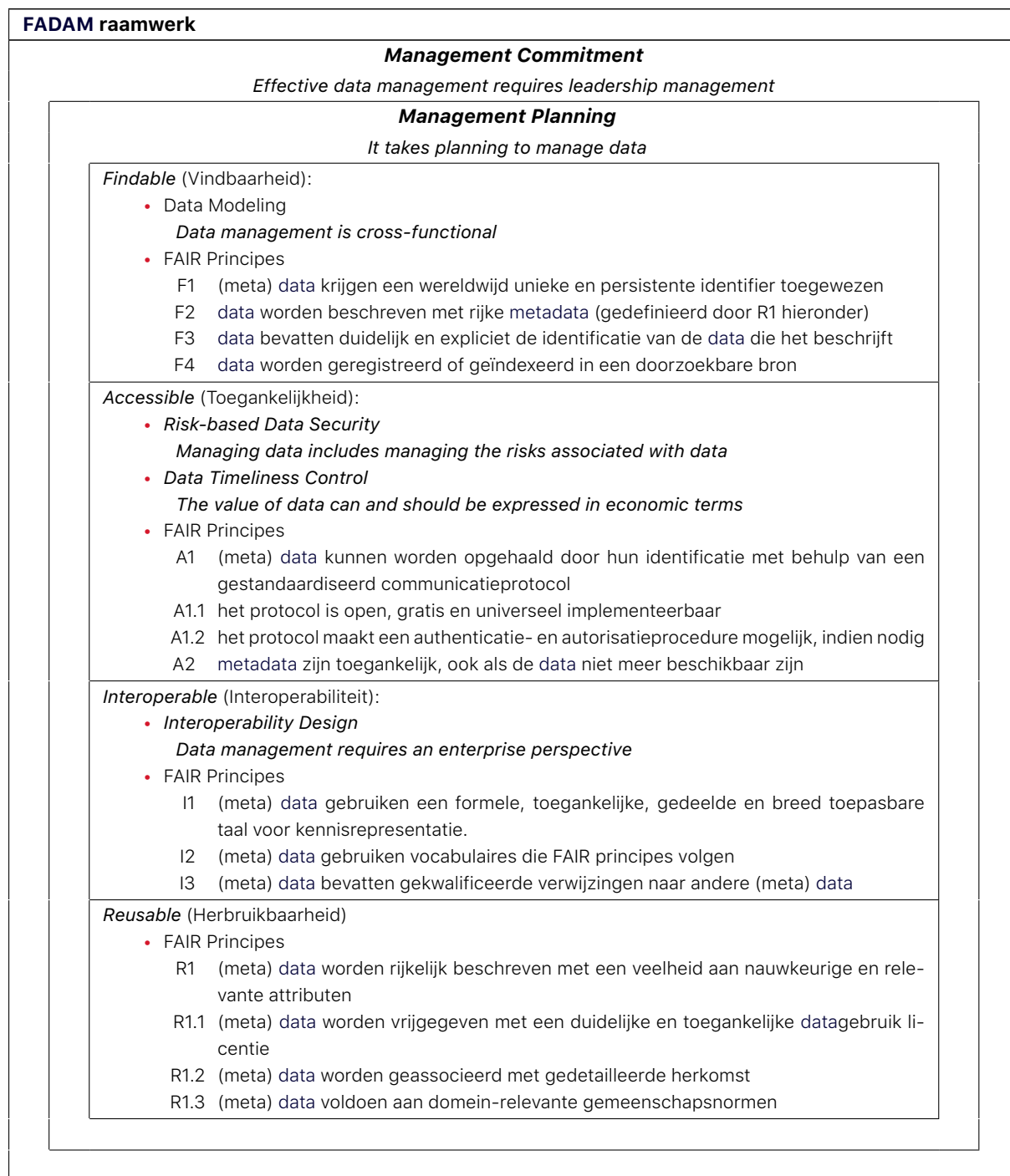
In tabel 'A.4-FAIR leidende principes' is het FAIR Data Management (FADAM) raamwerk aangevuld met de FAIR Data Principles om zo een compleet FAIR FADAM schema te krijgen.

FAIR leidende principes	
Vindbaar zijn:	
F1.	(meta) data krijgen een wereldwijd unieke en persistente identifier toegewezen
F2.	data worden beschreven met rijke metadata (gedefinieerd door R1 hieronder)
F3.	(meta) data bevatten duidelijk en expliciet de identificatie van de data die het beschrijft
F4.	(meta) data worden geregistreerd of geïndexeerd in een doorzoekbare bron
Toegankelijk zijn:	
A1.	(meta) data kunnen worden opgehaald door hun identificatie met behulp van een gestandaardiseerd communicatieprotocol
A1.1	het protocol is open, gratis en universeel implementeerbaar
A1.2	het protocol maakt een authenticatie- en autorisatieprocedure mogelijk, indien nodig
A2.	metadata zijn toegankelijk, ook als de data niet meer beschikbaar zijn
Interoperabel zijn:	
I1.	(meta) data gebruiken een formele, toegankelijke, gedeelde en breed toepasbare taal voor kennisrepresentatie.
I2.	(meta) data gebruiken vocabulaires die FAIR principes volgen
I3.	(meta) data bevatten gekwalificeerde verwijzingen naar andere (meta) data
Om herbruikbaar te zijn:	
R1.	(meta) data worden rijkelijk beschreven met een veelheid aan nauwkeurige en relevante attributen
R1.1	(meta) data worden vrijgegeven met een duidelijke en toegankelijke data gebruik licentie
R1.2	(meta) data worden geassocieerd met gedetailleerde herkomst
R1.3	(meta) data voldoen aan domein-relevante gemeenschapsnormen

Tabel A.4: FAIR leidende principes [Wilkinson, 2016]

FAIR Data Management

In tabel 'A.5-FADAM FAIR raamwerk' is het FADAM raamwerk aangevuld met de FAIR Data Principles om zo een compleet FAIR FADAM schema te krijgen.



Tabel A.5: FADAM FAIR raamwerk [dFigueiredo, 2019; Force11, 2020; Wilkinson, 2016]

A.2.4. Community Models

In deze paragraaf zijn de *Best Practices* Community Models (CM) beschreven die voor dit onderzoek zijn gebruikt.

Community Capability Model Framework

In tabel 'A.6-*Best Practices* Community Capability Model Framework (CCMF)' zijn de CCMF *Best Practices* van Lyon, Alexander Ball e.a. [Lyon, 2012b] beschreven.

Best Practices CCMF	
Collaboration	<ul style="list-style-type: none"> • Collaboration within the discipline/sector • Collaboration and interaction across disciplines • Collaboration and interaction across sectors • Collaboration with the public
Skills and training	<ul style="list-style-type: none"> • Skill sets • Pervasion of training
Openness	<ul style="list-style-type: none"> • Openness in the course of research • Openness of published literature • Openness of data • Openness of methodologies/workflows • Re-use of existing data
Technical infrastructure	<ul style="list-style-type: none"> • Computational tools and algorithms • Tool support for data capture and processing • Data storage • Support for curation and preservation • Data discovery and access • Integration and collaboration platforms • Visualisations and representations • Platforms for citizen science
Common practice	<ul style="list-style-type: none"> • Data formats • Data collection methods • Processing workflows • Data packaging and transfer protocols • Data description • Vocabularies, semantics, ontologies • Data identifiers • Stable, documented APIs
Economic and business model	<ul style="list-style-type: none"> • Sustainability/geographic scale/size of funding for research/infrastructure • Public-private partnerships • Productivity and return on investment
Legal and ethical issues	<ul style="list-style-type: none"> • Legal and regulatory frameworks • Management of ethical responsibilities and norms
Academic culture	<ul style="list-style-type: none"> • Entrepreneurship, innovation and risk • Reward models for researchers • Quality and validation frameworks

Tabel A.6: CCMF [Lyon, 2012b]

Community Driven Data Curation

In tabel 'A.7-Best Practices Community Driven Data Curation (CDDC)' zijn de CDDC *Best Practices* van Curry e.a. [Curry, 2010] beschreven.

Best Practices CDDC	
Social Best Practices	
Participation	Stakeholders involvement for both data producers and consumers must occur early in the project. This will help provide insight into the basic questions of what they want to do, for whom, and what it will provide. White papers are an effective means to present these ideas, and solicit opinion from the community. They can be used to establish an informal 'social contract' for the community.
Engagement	Outreach activities are essential for promotion and feedback. Social communication and networking forums are useful but be aware that with typical consumers-to-contributors ratios of less than 5%, the majority of your community may not communicate using these media. The communication by email still remains important.
Incentives	For a community to participate in sheer curation there must be a line of sight from the data curating activity, to tangible exploitation benefits. If the general community lacks awareness of the value proposition, a collaborative contribution environment will be slow to emerge. Recognizing contributing curators through a formal feedback mechanism will help reinforce the contribution culture and directly increase output quality.
Community Governance Models	An effective governance structure is vital to ensure the success of a community. Internal community's and consortium perform well when they leverage traditional corporate and democratic governance models. However, these approaches are not appropriate for open community's where there is a need to engage the community within the governance process. Successful governance models for open community's follow less orthodox approaches using <i>meritocratic</i> and <i>autocratic</i> principles.
Technical Best Practices	
Data Representation	Data representations that are robust and standardized will encourage community usage, and tools development. Support for legacy data formats should be considered, as is the ability to translate all data forward to deal with emergent technology and standards is important.
Balancing Human- and Computer-based Curation	Arriving at a balance between orchestrating automated and human assisted curation will improve data quality. For large datasets robotic curation should be used for validating data deposition and entry, while the community targets focused curation tasks. Robotic curation should always defer to, and never override, human curation edits.
Track Provenance	A user consuming data generated from third parties needs mechanisms to assess the entities and process involved in the generation and publication of this data. Provenance is a key aspect in the process of mapping the historical trail behind an information artifact and can help determining if the data is high quality, trustworthy and compliant. Different users can lead to different perspectives of provenance. A scientist may need to evaluate the fine grained experiment description behind the data, while for a business analyst the 'brand' of data provider can be sufficient for determining quality. The ability to provide a provenance description attached to the data plays an important role in the data quality process. All curation activities including edits, especially where human curators are involved should be recorded and maintained as part of a larger data provenance effort.
Data Consumption Infrastructure	As open datasets become more prevalent, companies will need to develop appropriate internal infrastructures to consume, curate, manage and integrate third-party data. External data can be generated by business partners, expert community's or from the open web, the organizations data governance policies will need to cater for this consumption.

Tabel A.7: CDDC [Curry, 2010]

A.2.5. Beantwoording literatuurvragen

Voor het beantwoorden van de literatuurvragen in paragraaf '2.3.6-Beantwoording literatuurvragen' in de onderstaande schematische samenvatting gemaakt.

In de linkerkolom van deze tabel is het gevonden wetenschappelijk onderzoeksveld weergegeven met de gevonden principes en richtlijnen. Als eerste wordt aangegeven welke literatuurvraag met de gevonden principes en richtlijnen beantwoord kan worden. Alleen de eerst keer wordt de volledige literatuurvraag weergegeven, daarna alleen een verwijzing. De principes en richtlijnen zijn gelabeld en genummerd.

Literatuurvraag	Dimensie	Verwijzing
L1. (Welke principes en richtlijnen van EA kunnen helpen om de sturing op het beleid van BD het hergebruik te verbeteren?)		
Enterprise Architecture Management		
EA principes:		
EA-1	Nemen van beslissingen	
EA-2	Coördinatie	
EA-3	Algemene acceptatie van EAM door de organisatie	§A.2.1
EA-4	Algemene kenmerken van de organisatiecultuur	
EA-5	Culturele verschillen tussen architecten en managers	
EA-6	Culturele verschillen tussen verschillende architecten	
L1. en L5. (Op welke manier kan gestuurd worden op het beleid rond BD?)		
Government Enterprise Architecture Framework		
EA / BD richtlijnen:		
EA-BD-1	Functionele geschiktheid	
EA-BD-2	Betrouwbaarheid	
EA-BD-3	Prestatie-efficiëntie	§A.2.1
EA-BD-4	Bruikbaarheid	
EA-BD-5	Beveiliging	
EA-BD-6	Compatibility	
L2. en L5. (Welke principes en richtlijnen van DG kunnen helpen om de sturing op het beleid van BD het hergebruik te verbeteren?)		
Open Government Data		
DG / BD principes:		
DG-BD-1	Visie	
DG-BD-2	Leiderschap	
DG-BD-3	Financiering Open data	§A.2.2
DG-BD-4	Zelf-ontwikkeld vermogen	
DG-BD-5	Communicatie	
L2. en L5. (Welke principes en richtlijnen van DG kunnen helpen om de sturing op het beleid van BD het hergebruik te verbeteren?)		
Master Data Management		
DG / BD richtlijnen:		
DG-BD-6	Interoperabiliteit	§A.2.2
DG-BD-7	Metadata	
L3. en L5. (Op welke manier kan het FADAM raamwerk helpen het hergebruik van BD efficiënter te maken?)		
FAIR Data Management		
FADAM richtlijnen:		
FADAM-BD-1	Management Commitment	
FADAM-BD-2	Managementplanning	
FADAM BD principes:		
FADAM-BD-3	Vindbaarheid	§A.2.3
FADAM-BD-4	Toegankelijkheid	
FADAM-BD-5	Interoperabiliteit	
FADAM-BD-6	Herbruikbaarheid	

Tabel A.8: Beantwoording literatuurvragen [dFigueiredo, 2019; Khatri, 2010; Kotusev, 2017; Lnenicka, 2019; Lyon, 2012b; Welle Donker, 2017; Wilkinson, 2016]

Literatuurvraag	Dimensie	Verwijzing
L4. (Welke <i>Best Practices</i> van CM helpen het hergebruik van BD breder en efficiënter te maken?)		
Community Capability Model Framework		
CM capaciteitsfactoren:		
	CM-BD-1	Samenwerking
	CM-BD-2	Vaardigheden en training
	CM-BD-3	Openheid
	CM-BD-4	Technische infrastructuur
	CM-BD-5	Gemeenschappelijke participatie
	CM-BD-6	Economisch en bedrijfsmodel
	CM-BD-7	Juridische en ethische kwesties
	CM-BD-8	Academische cultuur
		§A.2.4
L4. (Welke <i>Best Practices</i> van CM helpen het hergebruik van BD breder en efficiënter te maken?)		
Community Driven Data Curation		
CM capaciteitsfactoren:		
	CM-BD-9	Participatie
	CM-BD-10	Engagement
	CM-BD-11	Stimulansen
	CM-BD-12	Community Governance Models
		§A.2.4

Tabel A.8: Beantwoording literatuurvragen [dFigueiredo, 2019; Khatri, 2010; Kotusev, 2017; Lnenicka, 2019; Lyon, 2012b; Welle Donker, 2017; Wilkinson, 2016]

A.3. Methodologie

In deze paragraaf is de methodologie van hoofdstuk 3 weergegeven.

A.3.1. Conceptueel ontwerp

In tabel 'A.9-Kernbegrippen, dimensies en indicatoren' staan alle indicatoren beschreven op basis van de kernbegrippen van het conceptueel model van figuur '2.3-Conceptueel model'.

Variabele	kernbegrip	dimensie (principe, richtlijn, of Best Practices)	indicator
Onafhankelijke	EA principes [Kotusev, 2017]	EA-1: Nemen van beslissingen	<ul style="list-style-type: none"> Managers gebruiken EA artefacten voor het nemen van belangrijke beslissingen Architecten gebruiken EA artefacten voor het nemen van belangrijke beslissingen
		EA-2: Coördinatie	<ul style="list-style-type: none"> Samenwerking op ondernemingsniveau en businessunits. Lokale en globale behoeften in evenwicht brengen Gedeelde architectuurvisie
		EA-3: Algemene acceptatie van EAM door de organisatie	<ul style="list-style-type: none"> Organisatie is bekend met EAM EAM wordt geaccepteerd in de organisatie
		EA-4: Algemene kenmerken van de organisatiecultuur	<ul style="list-style-type: none"> Er zijn geen organisatie subculturen. Basis van waarheid Aard van de tijd Oriëntatie op het werk Oriëntatie op te veranderen Controle en verantwoordelijkheid Oriëntatie en focus Oriëntatie en Informatietechnologie (IT)
		EA-5: Culturele verschillen tussen architecten en managers	<ul style="list-style-type: none"> Er zijn geen organisatie subculturen tussen architecten en managers. Basis van waarheid tussen architecten en managers Aard van de tijd tussen architecten en managers Oriëntatie op het werk tussen architecten en managers Oriëntatie op te veranderen tussen architecten en managers Controle en verantwoordelijkheid tussen architecten en managers Oriëntatie en focus tussen architecten en managers Oriëntatie en IT tussen architecten en managers
		EA-6: Culturele verschillen tussen verschillende architecten	<ul style="list-style-type: none"> Er zijn geen organisatie subculturen tussen architecten onderling. Basis van waarheid tussen architecte onderling Aard van de tijd tussen architecten onderling Oriëntatie op het werk tussen architecten onderling Oriëntatie op te veranderen tussen architecten onderling Controle en verantwoordelijkheid tussen architecten onderling Oriëntatie en focus tussen architecten onderling Oriëntatie en IT tussen architecten onderling
Onafhankelijke	EA / BD richtlijnen [Lnenicka, 2019]	EA-BD-1: Functionele geschiktheid	<ul style="list-style-type: none"> Functionele compleetheid Functionele correctheid
		EA-BD-2: Betrouwbaarheid	<ul style="list-style-type: none"> Beschikbaarheid Fouttolerantie
		EA-BD-3: Prestatie-efficiëntie	<ul style="list-style-type: none"> Tijdgedrag
		EA-BD-4: Bruikbaarheid	<ul style="list-style-type: none"> Bescherming tegen gebruikersfouten
		EA-BD-5: Beveiliging	<ul style="list-style-type: none"> Vertrouwelijkheid Integriteit
		EA-BD-6: Compatibility	<ul style="list-style-type: none"> Interoperability

Tabel A.9: Kernbegrip, dimensies en indicatoren [Doorewaard, 2019, p. 62]

Variabele	Kernbegrip	dimensie (principe, richtlijn, of Best Practices)	indicator
Onafhankelijke	DG / BD principes [Welle Donker, 2017]	DG-BD-1: Visie	<ul style="list-style-type: none"> • Een gemeenschappelijk doel hebben • Gefragmenteerde aanpak vermijden • Samenwerking tussen belanghebbenden stimuleren
		DG-BD-2: Leiderschap	<ul style="list-style-type: none"> • Open data heeft een probleem eigenaar nodig • Open data heeft een eigenaar nodig die activiteiten stimuleert en coördineert • Bewustwording en capaciteitsopbouw • Politiek draagvlak opbouwen • Werkvloer ondersteuning
		DG-BD-3: Financiering Open data	<ul style="list-style-type: none"> • Zorgen voor voldoende financiële middelen op lange termijn • Budgetten voor gegevensbeheer en infrastructuren
		DG-BD-4: Zelf-ontwikkeland vermogen	<ul style="list-style-type: none"> • Vraag en aanbod aansluitend maken • Dataproviders proactief open data met vragen helpen • Dataproviders helpen bij vraag en aanbod
		DG-BD-5: Communicatie	<ul style="list-style-type: none"> • Weten met wie, hoe en er gecommuniceerd wordt. • Beginfase zal dit vooral interne communicatie • Latere fasen ook externe communicatie
Onafhankelijke	DG / BD richtlijnen [Khatri, 2010]	DG-BD-6: Interoperabiliteit	<ul style="list-style-type: none"> • Gebruik van data voor de business • Communicatie methoden voor data gebruik • Gedrag rond data ontplooiingen • Ontwikkelen van mogelijkheden data hergebruik • Reguleren invloed van de omgeving op het gebruik van data
		DG-BD-7: Metadata	<ul style="list-style-type: none"> • Programma voor het documenteren van semantiek voor data • Beschrijven van data consistentie en modellen zodat deze kunnen worden geïnterpreteerd • Plannen voor het <i>up-to-data</i> houden van metadata

Tabel A.9: Kernbegrip, dimensies en indicatoren [Doorewaard, 2019, p. 62]

Variabele	Kernbegrip	dimensie (principe, richtlijn, of Best Practices)	indicator
Onafhankelijke	FADAM richtlijnen [d'Figueiredo, 2019]	FADAM-BD-1: Management Commitment	<ul style="list-style-type: none"> • Effectief gegevensbeheer vereist leiderschap
		FADAM-BD-2: Managementplanning	<ul style="list-style-type: none"> • Er is planning nodig om gegevens te beheren • Het managing van data betekend het managing data kwaliteit
Onafhankelijke	FADAM principes [Force11, 2020; Wilkinson, 2016]	FADAM-BD-3: Vindbaarheid	<ul style="list-style-type: none"> • (Meta) data krijgen een wereldwijd unieke en persistente identifier toegewezen • Data worden beschreven met rijke metadata • (Meta) data bevatten duidelijk en expliciet de identificatie van de data die het beschrijft • (Meta) data worden geregistreerd of geïndexeerd in een doorzoekbare bron
		FADAM-BD-4: Toegankelijkheid	<ul style="list-style-type: none"> • (Meta) data kunnen worden opgehaald door hun identificatie met behulp van een gestandaardiseerd communicatieprotocol • Het protocol is open, gratis en universeel implementeerbaar • Het protocol maakt een authenticatie- en autorisatieprocedure mogelijk, indien nodig • Metadata zijn toegankelijk, ook als de data niet meer beschikbaar zijn
		FADAM-BD-5: Interoperabiliteit	<ul style="list-style-type: none"> • (Meta) data gebruiken een formele, toegankelijke, gedeelde en breed toepasbare taal voor kennisrepresentatie • (Meta) data gebruiken vocabulaires die FAIR principes volgen • (Meta) data bevatten gekwalificeerde verwijzingen naar andere (meta) data
		FADAM-BD-6: Herbruikbaarheid	<ul style="list-style-type: none"> • (Meta) data worden rijkelijk beschreven met een veelheid aan nauwkeurige en relevante attributen • (Meta) data worden vrijgegeven met een duidelijke en toegankelijke datagebruik licentie • (Meta) data worden geassocieerd met gedetailleerde herkomst • (Meta) data voldoen aan domein-relevante gemeenschapsnormen

Tabel A.9: Kernbegrip, dimensies en indicatoren [Doorewaard, 2019, p. 62]

Variabele	Kernbegrip	dimensie (principe, richtlijn, of Best Practices)	indicator
Modererende	CM capaciteitsfactoren [Lyon, 2012b]	CM-BD-1: Samenwerking CM-BD-2: Vaardigheden en training CM-BD-3: Openheid CM-BD-4: Technische infrastructuur CM-BD-5: Gemeenschappelijke participatie CM-BD-6: Economisch en bedrijfsmodel CM-BD-7: Juridische en ethische kwesties CM-BD-8: Academische cultuur	<ul style="list-style-type: none"> • Samenwerking binnen de discipline / sector • Samenwerking buiten de discipline / sector • Samenwerking / interactie met verschillende sectoren • Publieke samenwerking • Set met vaardigheden • Training programma's • Openheid in onderzoek • Openheid in gepubliceerde literatuur • Openheid in data • Openheid in methodologies en werkwijze • Hergebruik van beschikbare data • Rekenkundige tools en algoritmen • Tools voor aanmaken van data en het verwerken ervan • Dataopslag • Ondersteuning voor beheer en behoud • Gegevens ontdekken en toegang • Integratie- en samenwerkingsplatforms • Visualisatie en representatie van data • Publieke onderzoeksplatforms • Dataformaten • Methoden voor het verzamelen van data • Verwerken van <i>workflows</i> • <i>Datapackaging</i> en overdrachtsprotocollen • Beschrijven van data • Vocabulary, semantiek, ontologieën • Data identificaties • Stabiele en gedocumenteerde Application Programming Interface (API)'s • Financieringsmodellen voor onderzoek en infrastructuur • Publiek-private samenwerkingen • Wettelijke en regelgevende kaders • Beheer van ethische verantwoordelijkheden en normen • Productiviteit en rendement op investering • Ondernemerschap, innovatie en risico • Beloningsmodellen voor onderzoekers • Kwaliteitskaders en validatiekaders
Modererende	CM principes [Curry, 2010]	CM-BD-9: Participatie CM-BD-10: Engagement CM-BD-11: Stimulansen CM-BD-12: Community Governance Models	<ul style="list-style-type: none"> • Betrokkenheid van belanghebbenden voor zowel data producenten als consumenten • Outreach-activiteiten zijn onmisbaar voor promotie en feedback. • Sociale communicatie- en netwerkforums zijn nuttig • Exploitatievoordelen voor hergebruik data. • Een effectieve bestuursstructuur is onmisbaar om het succes van een gemeenschap te verzekeren • Interne gemeenschappen en consortium presteren goed wanneer ze gebruikmaken van traditionele bedrijfsmodellen en democratische bestuursmodellen

Tabel A.9: Kernbegrip, dimensies en indicatoren [Doorewaard, 2019, p. 62]

B

Vooronderzoek

Ter voorbereiding van de gebiedsverkenning van het hoofdonderzoek is er een vooronderzoek gedaan naar de mogelijke problemen die er zijn met de sturing op publieke overheidsdata binnen Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK). Het doel van dit vooronderzoek is het probleemgebied van het hoofdonderzoek te bepalen. Met als resultaat een probleemstelling om daarmee de vraagstelling van het hoofdonderzoek te helder te krijgen.

Voor dit vooronderzoek is zoveel mogelijk dezelfde structuur gebruikt als voor het hoofdonderzoek, met uitzondering van de paragrafen "Theoretisch kader", "Motivatie / relevantie" en "Aanpak in hoofdlijnen". In paragraaf 'B.1-Introductie' wordt kort een introductie van dit vooronderzoek gegeven en wordt de onderzoeks-aanpak vastgesteld. De resultaten van dit vooronderzoek zijn beschreven in paragraaf B.2 en wordt afgesloten met de discussie en conclusies in paragraaf B.3.

B.1. Introductie

De introductie voor dit vooronderzoek bestaan uit een korte uitleg van de achtergrond, paragraaf B.1.1 en een beperkte gebiedsverkenning in paragraaf 'B.1.2-Gebiedsverkenning'. Aansluitend de probleemstelling in paragraaf B.1.3 en de opdrachtformulering in paragraaf B.1.4. Deze introductie wordt afgesloten met de onderzoeks-aanpak in paragraaf B.1.5. De paragrafen voor motivatie en relevantie zijn hierin niet overgenomen. De motivatie en relevantie voor dit vooronderzoek is om de probleemstelling duidelijk te krijgen en welke problemen worden ondervonden met de sturing op het hergebruik van publieke overheidsdata binnen BZK.

B.1.1. Achtergrond

De Nederlandse overheid heeft een duidelijke taakstelling om gebruik van publieke overheidsdata te bevorderen en deze op een zo toegankelijk mogelijke manier beschikbaar te stellen aan burgers, bedrijven en instellingen. Voor vrij toegankelijke publieke gegevens wordt dit gedaan in de vorm van open data. Daarmee is de publieke overheidsdata te verdelen in open data en alles wat niet vrij toegankelijke is, dat definiëren wij als semi-open data, zie paragraaf 'A.1.1-Publieke overheidsdata'.

In 2019 heeft het BZK namens het kabinet de Data Agenda Overheid (DAO) gepubliceerd. De DAO is een interbestuurlijke agenda, waarin staat wat de overheid gaat doen om beter om te gaan met persoonsgegevens, open data en Big Data (BD). In deze agenda staat beschreven hoe data beter ten goede kunnen komen aan beleidsvorming en het oplossen van maatschappelijke vraagstukken door de overheid. Om een beter inzicht te krijgen in het gebruik van publieke overheidsdata heeft BZK een onderzoek laten uitvoeren door Donker e.a. [Donker, 2019].

De uitdagingen waar BZK voor staat is met name de sturing op het hergebruik van open data. Het doel van dit vooronderzoek is om duidelijk te krijgen welke problemen er worden gevoeld op het gebied van sturing op het hergebruik van publieke overheidsdata.

B.1.2. Gebiedsverkenning

Het onderzoek van Donker e.a. [Donker, 2019] is als uitgangspunt genomen van dit empirisch vooronderzoek. In het onderzoek van Donker e.a. [Donker, 2019] zijn antwoorden gegeven op de volgende onderzoeksvragen over het hergebruik van open data:

- Wie zijn open data hergebruikers?
- Op welke manier(en) gebruiken hergebruikers open data?
- Waar lopen hergebruikers tegenaan bij het gebruiken van open data?
- Wat zijn de behoeften van de hergebruikers?

Hergebruikers

Hergebruikers van publieke overheidsdata spelen ook in het hoofdonderzoek een belangrijke rol. Deze hergebruikers zijn op basis van de rol die een gebruiker van open data speelt onder te verdelen in: Verzamelaars, Enablers, Ontwikkelaars, Verrijkers en Eindgebruikers [Analytics, 2013; Donker, 2019]. Zie bijlage 'A-Verdieping', paragraaf 'A.1.2-Hergebruikers' voor meer informatie over deze verschillende hergebruikers.

Relevant onderzoek

In het onderzoek van Donker e.a. [Donker, 2019] over het hergebruik van open data zijn een aantal bevindingen gedaan. Op basis van deze bevindingen zijn er een aantal aanbevelingen gedaan.

Hieronder een selectie van de belangrijkste bevindingen die ook van belang zijn in dit vooronderzoek [Donker, 2019]:

- Uit dit onderzoek blijkt dat open data op grote schaal en door een diverse groep worden gebruikt voor allerlei doeleinden en via allerlei wegen.
- Data.overheid.nl lijkt vooral gebruikt te worden als een catalogusdienst voor andere dataplatforms.
- Uit dit onderzoek blijkt dat hergebruikers vooral belemmeringen ervaren met open data van lokale overheden. Het aanbod is versnipperd, er is geen eenduidige standaardisatie met betrekking tot de datastructuur en bestandsformaten, en niet alle lokale overheden maken gelijksoortige datasets beschikbaar als open data waardoor er geen landdekking of regiodekking bereikt kan worden.
- Uit dit onderzoek blijkt ook dat hergebruikers vooral behoefte hebben aan meer samenwerking op het gebied van open data. Hergebruikers willen graag betrokken worden bij de governance van open data proces.

In het onderzoek van Donker e.a. [Donker, 2019] zijn een aantal aanbevelingen gedaan aan BZK. Deze aanbevelingen zijn zowel belangrijk voor dit vooronderzoek, maar ook voor het hoofdonderzoek. Hieronder een selectie van de belangrijkste aanbevelingen [Donker, 2019]:

- Ontwikkel geautomatiseerde processen als onderdeel van een Data Governance (DG) strategie om data sneller te ontsluiten als open data.
- Data via verplichte standaarden publiceren, inclusief gestandaardiseerde en consistente codering voor de structuur.
- Het zou echter de governance van open data ten goede komen als de overheid alle overheidsdata – ook de niet openbare data – via een centraal datapunt zouden gebruiken.
- Om datagedreven te kunnen werken moeten data kunnen stromen. En dat kan alleen als er een datastrategie belegd is zowel op macroniveau (nationaal) als op mesoniveau (regionaal) en microniveau (organisatie).
- Datakwaliteit wordt ook steeds belangrijker naarmate meer geautomatiseerde besluitvorming gebaseerd is op algoritmes. En dat vergt dat overheden transparant zijn over de bron en de inhoud van de data. Daarom is het essentieel dat er een duidelijke lange termijn visie aanwezig is en dat de geplande fases ook werkelijkheid worden.

B.1.3. Probleemstelling

De Nederlandse rijksoverheid en de decentrale overheid voeren gezamenlijk de DAO uit. Met DAO geeft het BZK invulling aan de ambitie van de overheid: de mogelijkheden van Data optimaal benutten met oog voor publieke waarden [Digitale Overheid, 2020a]. Een van de actiepunten is het bevorderen van hergebruik van publieke overheidsdata. Hiervoor is het onderzoek van Donker e.a. [Donker, 2019] uitgevoerd en zijn er een aantal bevindingen en aanbevelingen gedaan, zoals beschreven in paragraaf 'B.1.2-Gebiedsverkenning'.

In het onderzoek van Donker e.a. [Donker, 2019] zijn een aantal bevindingen en aanbevelingen gedaan aan BZK. Zoals het ontwikkelen van geautomatiseerde processen als onderdeel van de DG strategie, het verplichten van publicatiestandaarden, en dat er een centraal datapunt komt voor alle overheidsdata.

De probleemstelling van dit vooronderzoek is:

- *Worden de bevindingen en aanbevelingen van Donker e.a. [Donker, 2019] ook door BZK onderkend?*

Het probleemgebied van zowel dit vooronderzoek als het hoofdonderzoek moet binnen het onderzoeksveld liggen van BD en Enterprise Architectuur (EA).

B.1.4. Opdrachtformulering

Binnen BZK is directie Digitale Overheid (DO) verantwoordelijk voor het maken van beleid rond de DAO en is de probleemeigenaar van het verbeteren van het hergebruik van open data, daarmee de opdrachtgever binnen dit onderzoek. Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) zorgt binnen BZK voor de publicatie van alle officiële informatie van de Nederlandse (rijks)overheid, waaronder open data. Het doel van dit vooronderzoek is, te achterhalen welke problemen er zijn met de sturing op het hergebruik van open data die onder de verantwoording van BZK valen.

Op basis van de probleemstelling is de volgende hoofdvraag te formuleren voor het interview:

- *Welke problemen zijn er op het gebied van sturing op het hergebruik van open data?*

B.1.5. Onderzoeksaanpak

Als onderzoeksmethode voor dit vooronderzoek is gekozen om een interview te houden met twee coördinerende beleidsmedewerkers van DO. Met als doel inzicht te krijgen in de problemen die er binnen DO spelen met betrekking tot de sturing op het hergebruik van open data. Op basis van Saunders e.a. [Saunders, 2019, p.436-444] is gekozen om hiervoor een semigestructureerde interview te houden om zo dit probleemgebied te verkennen.

Als afbakening binnen het Business Process Management (BPM) domein, is gekozen voor een probleem binnen het onderzoeksveld van EA en DG.

Deze kwalitatieve data wordt vervolgens omgezet middels kwalitatief transcriberen naar tekst [Saunders, 2019, p.644-647] en gevalideerd door de geïnterviewden. Vervolgens wordt van deze kwalitatief tekst een samenvatting gemaakt [Saunders, 2019, p.649]. Daarna kan deze samengevatte kwalitatief tekst verder worden geanalyseerd door of thematische analyse [Saunders, 2019, p.651-660] toe te passen of door de geaarde theorie methode toe te passen [Saunders, 2019, p.668-674].

B.2. Resultaat

Op basis van de probleemstelling zijn er een aantal interview vragen geformuleerd. Deze vormen de basis van de semigestructureerde interview van paragraaf 'B.2.1-Interview'.

Daarna is er een volledige transcriptie gemaakt en deze is gevalideerd door de coördinerende beleidsmedewerkers van DO. Van deze volledige transcriptie is er een samenvatting gemaakt in paragraaf 'B.2.2-Samenvatting interview'.

Als laatste zijn de gegevens van het interview geanalyseerd en geëvalueerd in paragraaf 'B.2.3-Analyse en evaluatie'.

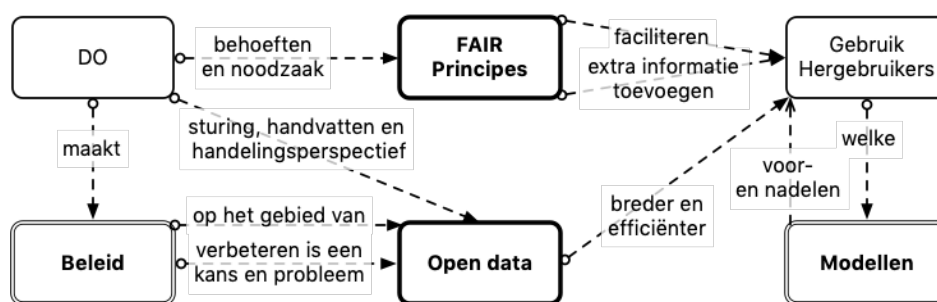
B.2.1. Interview

In een semigestructureerde interview van maximaal 30 minuten moet er een duidelijk probleemgebied worden gevonden die betrekking heeft op het onderzoeksveld van BD en EA. Hiervoor wordt het Situatiebeoordeling (SB) denkproces van Kepner en Tregoe [Kepner, 1997, Chap:Seven] gebruikt om verschillende probleemgebieden te vinden en te prioriteren. De reden dat er gekozen is voor het SB denkproces, is dat deze op een gestructureerde manier belangrijke aandachtspunten verzameld en op een gestructureerde manier prioriteerd.

Gezien de beperkte tijd is er voor gekozen om alle drie stappen van het SB denkproces tegelijkertijd uit te voeren. Tijdens dit interview zijn aandachtspunten verzameld met betrekking tot de sturing op het hergebruik van open data. Hierbij wordt gelijk gecontroleerd of deze voldoen aan de voorwaarden van stap 3 van de SB. Tegelijkertijd wordt ook de prioriteit bepaald als er meerdere aandachtspunten naar voren komen, dit is stap 2 van de SB.

B.2.2. Samenvatting interview

Van de samenvatte transcriptie zijn de stukken tekst in losse statements gezet. Deze losse statements dienen als basis voor thematische analyse van Saunders e.a. [Saunders, 2019, p.651-660].



Figuur B.1: Interview DO

In eerste instantie zijn voor de woorden *gebruikers*, *gebruikersgroep*, *doelgroepen* en *hergebruikers* is de term **Hergebruikers** gekozen. Voor de woorden *hergebruik* en *gebruik* is de term **Gebruik** gekozen. En voor de woorden *publieke overheidsdata* en *open data* is de term **Open data** gekozen.

Vervolgens is van deze losse statements een grafische weergave gemaakt. Bij de analyse van de grafische weergave kwamen de volgende vier kernbegrippen naar voren; **FAIR Data Principles**, **Open data**, **Models** en **Beleid**. Daarna is deze grafische weergave vereenvoudigd rond deze vier kernbegrippen waardoor figuur 'B.1-Interview DO' is ontstaan.

Hieronder de gevonden relaties tussen de verschillende entiteiten en de vier kernbegrippen. Deze relaties kunnen worden geëvalueerd om een antwoord te geven op de vragen van dit vooronderzoek.

- Kans en probleem om **Open data** **Beleid** te verbeteren.
- **Open data** kan breder en efficiënter worden **Gebruikt**.
- Het **Gebruik** van **Open data** verder faciliteren.
- Zijn de **FAIR Data Principles** bruikbaar het **Gebruik Open data** te faciliteren.
- **Gebruik** van **Open data** verbeteren door extra informatie toe te voegen.
- Kunnen **FAIR Data Principles** faciliteren **Open data** te voorzien van extra informatie.
- DO zoekt handvatten, handelingsperspectief en sturing **Open data**.
- Behoeften en noodzaak van principes zoals de **FAIR Data Principles** voor DO.
- Welke Data en community **Models**, kunnen worden toegepast op **Hergebruikers** en wat zijn de voor- en nadelen.
- DO maakt het **Beleid** op het gebied van **Open data**.
- DO zit op een hoger abstractie niveau.

B.2.3. Analyse en evaluatie

Op basis van de relaties tussen de verschillende entiteiten en de vier kernbegrippen kan het volgende antwoord worden gegeven op de hoofdvraag van dit vooronderzoek.

- *Welke problemen zijn er op het gebied van sturing op het hergebruik van open data?*

DO maakt het beleid op het gebied van open data en ziet dit als een kans en probleem om het open data beleid te verbeteren. Daarvoor heeft DO handvatten, handelingsperspectief en sturing nodig. DO heeft behoefte en noodzaak aan principes zoals FAIR Data Principles. Vragen die hierbij naar voren zijn komen:

- Kan het hergebruik van open data worden verbeterd door deze te voorzien van extra informatie en kunnen de FAIR Data Principles hierbij faciliteren?
- Welke data en Community Models kunnen worden toegepast om het hergebruik van open data breder en efficiënter te maken en wat zijn hierbij de voor- en nadelen?

B.3. Discussie en conclusie

B.3.1. Discussie

Het SB denkproces van Kepner en Tregoe [Kepner, 1997, Chap:Seven] is een goede methode om op een gestructureerde manier belangrijke aandachtspunten te verzamelen en deze op een gestructureerde manier te prioriteren. Dit denkproces wordt normaal gesproken niet gebruikt in combinatie met een semigestructureerde interview, vanwege de vaak beperkte interview tijd. Daarom zijn alle deelstappen van het SB denkproces in één keer uitgevoerd om te komen tot de belangrijkste probleemstelling.

B.3.2. Conclusies

Het doel van dit vooronderzoek is het probleemgebied van het hoofdonderzoek te bepalen en een aanzet geven tot een probleemstelling. Er is duidelijke behoefte aan sturing op het hergebruik van open data, dit wordt gezien als kans en probleem om het open data beleid te verbeteren. Het hergebruik van open data kan mogelijk worden verbeterd door deze van extra informatie te voorzien door het toepassen van de FAIR Data Principles. Daarnaast kan hergebruik van open data mogelijk breder en efficiënter door het toepassen van data en Community Models.

Kort samengevat laat de probleemstelling van dit afstudeeronderzoek zich omschrijven als volgt:

- *Er is een duidelijke behoefte aan sturing op het beleid van open data om het hergebruik te verbeteren. Het is onbekend of de FAIR Data Principles hierbij kunnen helpen door open data te voorzien van extra informatie. En welke data en Community Models kunnen worden toegepast om het hergebruik van open data breder en efficiënter te maken.*

C

Templates

C.1. Interview template

Hieronder de interview template waarin per kernbegrip (Enterprise Architectuur (EA), EA–Data Governance (DG), FAIR Data Management (FADAM)–Big Data (BD) en Community Models (CM)–BD) de interviewvragen worden gepresenteerd. Niet alle respondenten krijgen dezelfde vragen, zie hiervoor tabel '4.1-Respondenten en onderzoeksvragen' welke vragen per functie/taak worden gesteld.

De vragen worden gegroepeerd per dimensie en per dimensie wordt na het beantwoorden van de vraag een score gevraagd te geven op basis van de aangegeven scorematrix.

C.1.1. Samenvatting transcriptie interview

Op basis van tabel 'A.9-Kernbegrippen, dimensies en indicatoren' waarin de kernbegrippen en de dimensies zijn weergegeven zijn de onderstaande interviewvragen samengesteld. Een topic is een samenvoeging van een dimensie en het kernbegrip zoals de in tabel 'A.8-Beantwoording literatuurvragen' zijn aangegeven. Hiermee kunnen de literatuurvragen wordt beantwoord zoals is aangegeven in paragraaf '2.3.6-Beantwoording literatuurvragen'. In paragraaf 'A.2-Theoretische kader' zijn de verschillende dimensies verder uitgewerkt.

Interview	<Datum>
Rol	<Rol>
Organisatieonderdeel	<Organisatieonderdeel>
Positie in de organisatie	<Positie in de organisatie>
Plaats interview	<Online, via WebEx Meetings, fysieke locatie>
Tijd	<Begin en eind tijd en er was voldoende tijd? Totaal 0 minuten>
Jaren ervaring	<Jaren ervaring in deze rol>
Algemene indruk	<Algemene indruk van de geïnterviewde en een indicatie van de kwaliteit van de kwalitatieve data.>
Topic	Vraag
Toepassen van EA principes en richtlijnen	
EA-1 Nemen van beslissingen	Gebruiken managers EA artefacten voor het nemen van belangrijke beslissingen? Gebruiken architecten EA artefacten voor het nemen van belangrijke beslissingen?
EA-2 Coördinatie	Werken architecten op ondernemingsniveau en businessunits samen en waaruit blijkt dat? Brengeen architecten op ondernemingsniveau en businessunits lokale en globale behoeften in evenwicht en hoe wordt dit gedaan? Is er een gedeelde architectuurvisie bij architecten op ondernemingsniveau en businessunits en hoe wordt deze gedeeld?

Tabel C.1: Template transcriptie interview

Topic	Vraag
EA-3 Algemene acceptatie van EAM door de organisatie	Is de organisatie bekend met Enterprise Architecture Management (EAM) en hoe is deze geïmplementeerd? Wordt EAM geaccepteerd in de organisatie en waaruit blijkt dat?
EA-4 Algemene kenmerken van de organisatiecultuur	Zijn er verschillende subculturen binnen het domein van EA? Er zijn een aantal algemene kenmerken te noemen van de organisatiecultuur, zoals basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en Informatietechnologie (IT)?
EA-5 Culturele verschillen tussen architecten en managers	Zijn er verschillende subculturen tussen architecten en managers binnen het domein van EA? Zijn er culturele verschillen tussen architecten en managers op basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en IT?
EA-6 Culturele verschillen tussen verschillende architecten	Zijn er verschillende subculturen tussen architecten onderling binnen het domein van EA? Zijn er culturele verschillen tussen verschillende architecten op basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en IT?
EA-1 t/m 6 Invloed op herbruikbaarheid publieke overheidsdata	Scorematrix huidige situatie: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk Scorematrix toepasbaarheid: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk
Toepassen van Enterprise Architectuur principes en richtlijnen rond Big Data	
EA-BD-1 Functionele geschiktheid	Welke regels of richtlijnen zijn er op het gebied functionele compleetheid met betrekking tot publieke overheidsdata? Welke regels of richtlijnen zijn er op het gebied functionele correctheid met betrekking tot publieke overheidsdata?
EA-BD-2 Betrouwbaarheid	Welke regels of richtlijnen zijn er op het gebied beschikbaarheid met betrekking tot publieke overheidsdata? Welke regels of richtlijnen zijn er op het gebied fouttolerantie met betrekking tot publieke overheidsdata?
EA-BD-3 Prestatie-efficiëntie	Welke regels of richtlijnen zijn er op het gebied tijdgedrag met betrekking tot publieke overheidsdata?
EA-BD-4 Bruikbaarheid	Welke regels of richtlijnen zijn er op het gebied bescherming tegen gebruikersfouten met betrekking tot publieke overheidsdata?
EA-BD-5 Beveiliging	Welke regels of richtlijnen zijn er op het gebied vertrouwelijkheid met betrekking tot publieke overheidsdata? Welke regels of richtlijnen zijn er op het gebied integriteit met betrekking tot publieke overheidsdata?
EA-BD-6 Compatibility	Welke regels of richtlijnen zijn er op het gebied interoperabiliteit met betrekking tot publieke overheidsdata?
EA-BD-1 t/m 6 Invloed op herbruikbaarheid publieke overheidsdata	Scorematrix huidige situatie: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk Scorematrix toepasbaarheid: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk

Tabel C.1: Template transcriptie interview

Topic	Vraag
Toepassen van Data Governance principes en richtlijnen rond Big Data	
DG-BD-1 Visie	Wat is de visie met betrekking tot het hergebruik van publieke overheidsdata?
DG-BD-2 Leiderschap	Welke vormen van leiderschap hebben betrekking tot het hergebruik van publieke overheidsdata?
DG-BD-3 Financiering Open data	Op welke manier wordt de financiering van publieke overheidsdata geregeld?
DG-BD-4 Zelf-ontwikkeland vermogen	Op welke manier wordt het zelf-ontwikkeland vermogen van het hergebruik van publieke overheidsdata vormgegeven?
DG-BD-5 Communicatie	Op welke manier wordt communicatie rond het hergebruik van publieke overheidsdata vormgegeven?
DG-BD-6 Interoperabiliteit	Welke data principes worden er toegepast om het hergebruik van publieke overheidsdata te verbeteren?
DG-BD-7 Metadata	Welke rol speelt metadata een rol bij het hergebruik van publieke overheidsdata?
DG-BD-1 t/m 7 Invloed op herbruikbaarheid publieke overheidsdata	Scorematrix huidige situatie: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk Scorematrix toepasbaarheid: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk
Toepassen van FAIR data management principes en richtlijnen rond Big Data	
FADAM-BD-1 Management Commitment	Op welke wijze speelt management commitment een rol bij het hergebruik van publieke overheidsdata?
FADAM-BD-2 Managementplanning	Op welke wijze speelt managementplanning een rol bij het hergebruik van publieke overheidsdata?
FADAM-BD-3 Vindbaarheid	Op welke wijze wordt de vindbaarheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?
FADAM-BD-4 Toegankelijkheid	Op welke wijze wordt de toegankelijkheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?
FADAM-BD-5 Interoperabiliteit	Op welke wijze wordt de interoperabiliteit van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?
FADAM-BD-6 Herbruikbaarheid	Op welke wijze wordt de herbruikbaarheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?
FADAM-BD-1 t/m 6 Invloed op herbruikbaarheid publieke overheidsdata	Scorematrix huidige situatie: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk Scorematrix toepasbaarheid: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk
Ontwikkelen van Best Practices Community Models rond Big Data	
CM-BD-1 Samenwerking	Hoe verloopt de samenwerking met de verschillende datacommunities?

Tabel C.1: Template transcriptie interview

Topic	Vraag
CM-BD-2 Vaardigheden en training	Wordt er bepaalde vaardigheden verwacht van de verschillende datacommunities en zijn daarvoor trainingen ingericht?
CM-BD-3 Openheid	Welke mate van openheid richting de verschillende datacommunities wordt er nagestreefd?
CM-BD-4 Technische infrastructuur	Worden er technische infrastructuur beschikbaar gesteld aan de verschillende datacommunities?
CM-BD-5 Gemeenschappelijke participatie	Hoe wordt er aandacht besteed aan gemeenschappelijke participatie met de verschillende datacommunities?
CM-BD-6 Economisch en bedrijfsmodel	Welke mate worden de verschillende datacommunities ondersteund op het gebied van economisch en bedrijfsmodellen?
CM-BD-7 Juridische en ethische kwesties	Hoe wordt met juridische en ethische kwesties omgegaan met de verschillende datacommunities?
CM-BD-8 Academische cultuur	Hoe is de academische cultuur vormgegeven bij de verschillende datacommunities?
CM-BD-9 Participatie	Hoe wordt de participatie met de verschillende datacommunities vormgegeven?
CM-BD-10 Engagement	Op welke wijze wordt engagement bewerkstelligd bij de verschillende communities?
CM-BD-11 Stimulansen	Op welke wijze wordt de verschillende communities gestimuleerd voor data hergebruik?
CM-BD-12 Community Governance Models	Worden door de verschillende datacommunities Governance models gebruikt?
CM-BD-1 t/m 12 Invloed op herbruikbaarheid publieke overheidsdata	Scorematrix huidige situatie: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk Scorematrix toepasbaarheid: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk
Overige, bijvangst	
Overige	<i>Zijn er nog andere interessante zaken?</i>
Bijvangst	<Onderwerp>

Tabel C.1: Template transcriptie interview

C.2. Template analyse documentatie

Per geanalyseerd document wordt de metadata vastgelegd met als doel deze te identificeren als gebruikte bron. Verder wordt hierin kort vastgelegd, het aantal citaten, de authenticatie en relevantie van het document voor dit onderzoek, maar ook de relatie met de onderzoeksvragen.

Document	<Naam (referentie)>
Titel	<titel>
Ontvangst informatie	<Organisatie, functie, tijdstip >
Auteur	<Auteur>
Publicatiedatum	<Publicatiedatum>
Status	<Status>
Versie	<Versie>
Taal	<Taal>
Beschikbaarheid	<Beschikbaarheid>
Aantal citaten	
Score document	
Totaal score document	
Gemiddelde score authenticatie	
Gemiddelde score datum publicatie	
Gemiddelde score relevantie	
Gemiddelde score scope	
Relatie document met de onderzoeksvragen	
Hoe wordt het document gebruikt door de actoren	
Samenvatting:	

Tabel C.2: Template analyse document

D

Kwalitatieve interview data

In deze bijlage staan de kwalitatieve interview data, bestaande uit de interview data en de likertschaal resultaten. In paragraaf 'D.2-Resultaten triangulatie antwoorden' is een samenvatting gemaakt, per dimensie en kernbegrip, van de transcripties van de interviews met de verschillende medewerkers van Nederlandse ministerie van Binnenlandse Zaken en Koninkrijksrelaties (BZK).

D.1. Samenvatting transcripties

D.1.1. Enterprise Architectuur

De interviewvragen over Enterprise Architectuur (EA) zijn beantwoord door de volgende respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker twee
- Datamanager
- Product Owner DONL

Nemen van beslissingen (EA-1)

Gebruiken managers EA artefacten voor het nemen van belangrijke beslissingen?

Gebruiken architecten EA artefacten voor het nemen van belangrijke beslissingen?

Het is niet duidelijk of managers EA artefacten bij het nemen van belangrijke beslissingen, hoewel enterprise architecten dit wel doen. Informatietechnologie (IT)-projecten mislukken niet omdat de enterprise architect het onvoldoende bedacht heeft, maar omdat het management niet in staat is om de gevolgen hiervan om te zetten in maatregelen binnen de organisatie. Men moet vooruitkijken en bouwen voor de toekomst, maar dit wordt niet altijd in acht genomen.

Coördinatie (EA-2)

Werken architecten op ondernemingsniveau en businessunits samen en waaruit blijkt dat?

Brengen architecten op ondernemingsniveau en businessunits lokale en globale behoeften in evenwicht en hoe wordt dit gedaan?

Is er een gedeelde architectuurvisie bij architecten op ondernemingsniveau en businessunits en hoe wordt deze gedeeld?

Architecten op ondernemingsniveau en businessunits werken samen om lokale en globale behoeften in balans te brengen. Er wordt gewerkt aan een gedeelde architectuur visie op ondernemingsniveau en businessunits, maar er is nog ruimte voor verbeteringen in de afstemming tussen deze organisaties. Er zijn nog steeds afzonderlijke eilandjes binnen de organisatie. Er is ambitie om beter samen te werken voor efficiëntie, maar er zijn nog steeds eilandjes. Binnen Kennis- en Exploitatiecentrum voor Officiële Overheidspublicaties (KOOP) is er een architectenteam, dat voor voornamelijk uit ontwerpers bestaat, maar niet elke groep heeft een eigen Enterprise architect. Waar architectuurprincipes zijn, wordt er geprobeerd om een architect onderdeel te laten

zijn van het team. Wat breed wordt gedragen is *loosely coupled systems* om Linked Data (LD) met elkaar te verbinden.

Algemene acceptatie van EAM door de organisatie (EA-3)

Is de organisatie bekend met Enterprise Architecture Management (EAM) en hoe is deze geïmplementeerd?

Wordt EAM geaccepteerd in de organisatie en waaruit blijkt dat?

De organisatie is op de hoogte van EAM, maar er is nog geen formele invulling. Er zit een architect in het Managementteam (MT) van KOOP. KOOP is momenteel in de fase van transformatie naar een volgende stap in volwassenheid, maar de details hiervan zijn nog niet bekend. Directie Digitale Overheid (DO) staat iets verder van de data, zij beheren geen data zelf, maar maken beleid voor organisaties die wel data beheren. DO werkt aan een federatief datastelsel en aan de ontwikkeling van een data architectuur die verschillende EA's beter met elkaar laat samenwerken. Dit kan bijvoorbeeld door gebruikt te maken van de Data Catalog Vocabulary (DCAT) standaarden.

Algemene kenmerken van de organisatiecultuur (EA-4)

Zijn er verschillende subculturen binnen het domein van EA?

Er zijn een aantal algemene kenmerken te noemen van de organisatiecultuur, zoals basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en IT?

Binnen KOOP bestaan er geen verschillende subculturen binnen het domein van EA. De architecten begrijpen elkaar als ook als ze niet altijd eens zijn. Er zijn duidelijke opdrachten en er is ruimte voor creativiteit. Als er ondersteuning nodig is, is deze beschikbaar, maar het is niet gebruikelijk dat overal een enterprise architect werkzaam is. Het coronavirus kan hierin ook een rol hebben gespeeld, binnen KOOP vindt veel informele besluitvorming plaats.

Culturele verschillen tussen architecten en managers (EA-5)

Zijn er verschillende subculturen tussen architecten en managers binnen het domein van EA?

Zijn er culturele verschillen tussen architecten en managers op basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en IT?

Er zijn geen grote cultuurverschillen tussen architecten en managers binnen het domein van EA. Architecten kunnen elkaar goed verstaan. Enterprise architectuur is een professionele functie dit meer benadrukt verdient omdat deze nuttig is. Er ontstaat vaak een brede maatschappelijke discussie, uiteindelijk wordt er geen besluit genomen en deze blijft doorschuiven.

Culturele verschillen tussen verschillende architecten (EA-6)

Zijn er verschillende subculturen tussen architecten onderling binnen het domein van EA?

Zijn er culturele verschillen tussen verschillende architecten op basis van waarheid, controle en verantwoordelijkheid, oriëntatie op het werk, veranderingen, focus en IT?

Er kan weinig gezegd worden over verschillende subculturen tussen architecten binnen het domein van EA, omdat er onvoldoende inzicht is vanuit de perspectieven van de DO en KOOP. Hierdoor is de vraag moeilijk te beantwoorden.

D.1.2. Enterprise Architectuur en Big Data

De interviewvragen over EA en Big Data (BD) zijn beantwoord door de volgende respondenten:

- Datamanager
- Product Owner DONL
- Adviseur Content Standaarden

Functionele geschiktheid (EA-BD-1)

Welke regels of richtlijnen zijn er op het gebied functionele compleetheid met betrekking tot publieke overheidsdata?

Welke regels of richtlijnen zijn er op het gebied functionele correctheid met betrekking tot publieke overheidsdata?

De functionele specificatie zijn wellicht duidelijk, maar de architectuur achter de meeste data is minder duidelijk. Er bestaat vaak een wettelijke basis voor datasets, die beschrijft waarom data openbaar of gepubliceerd moeten worden, maar niet hoe dit gedaan moet worden. Veel datasets die openbaar moeten worden gemaakt, worden vanuit eigen beweging door gemeenten of ministeries of provincies gepubliceerd vanuit data.overheid.nl, maar er is geen wettelijke grondslag behalve de Wet open overheid (Woo). De Woo en Wet hergebruik van overheidsinformatie (Who) worden steeds meer verplicht, maar in de meeste gevallen wordt er vaak heel weinig gespecificeerd. Veel zaken die onder BD vallen, zijn niet omschreven. Het beoordelen van de kwaliteit, volledigheid en integriteit van data blijkt een uitdaging te zijn die data.overheid.nl ondervindt, waarbij het beoordelen van de kwaliteitsaspecten een blijvende uitdaging is.

Betrouwbaarheid (EA-BD-2)

Welke regels of richtlijnen zijn er op het gebied beschikbaarheid met betrekking tot publieke overheidsdata?

Welke regels of richtlijnen zijn er op het gebied fouttolerantie met betrekking tot publieke overheidsdata?

Voor data.overheid.nl geldt dat de beschikbaarheid en fouttolerantie zeer beperkt zijn en er in de basis geen regels voor bestaan. Hier kom je ook in aanraking met de Woo en Who en de Uitvoeringswet Algemene Verordening Gegevensbescherming (AVG), waarbij juristen kanttekeningen plaatsen bij het delen van data omdat het ook AVG-proof moet zijn. Gebruikers kunnen feedback geven over een dataset op data.overheid.nl, maar de verbeteringen moeten door de aanbieder worden uitgevoerd. Kwaliteitscriteria of een garantie voor volledigheid, juistheid en accuraatheid zouden hierbij kunnen helpen, maar zo zit het spel nog niet op de wagen.

Prestatie-efficiëntie (EA-BD-3)

Welke regels of richtlijnen zijn er op het gebied tijdgedrag met betrekking tot publieke overheidsdata?

Voor data.overheid.nl gelden beperkte regels. Er is ruimte om zelf actie te ondernemen, maar het is belangrijk om eerst na te gaan hoe het is ingekaderd en welke regels er gehandhaafd moeten worden. Er wordt hard aan gewerkt om dit te verbeteren, maar de regels zijn nog niet uitgebreid vastgelegd. Bij sommige overheidsinstanties is het bijwerken van data een geautomatiseerd proces en maken ze gebruik van de data.overheid.nl Application Programming Interface (API)'s.

Bruikbaarheid (EA-BD-4)

Welke regels of richtlijnen zijn er op het gebied bescherming tegen gebruikersfouten met betrekking tot publieke overheidsdata?

Data.overheid.nl heeft onvoldoende inzicht in hoe de data daadwerkelijk wordt gebruikt. Dit is echter wel een belangrijk punt waarmee rekening wordt gehouden. Er wordt contact gelegd met gebruikers van data.overheid.nl om succesverhalen te delen. Aan de andere kant worden er ook dataverzoeken ontvangen waarmee problemen worden gemeld, wat inzicht geeft in knelpunten. Er is een opdracht uit Europa om nauwer contact te onderhouden met de gebruikers van de data. Op politiek niveau is er in de Tweede Kamer een nieuw bewindspersoon die hier actief mee bezig gaat houden. Er zijn eerder data-initiatieven ondernomen, maar deze waren niet succesvol.

Beveiliging (EA-BD-5)

Welke regels of richtlijnen zijn er op het gebied vertrouwelijkheid met betrekking tot publieke overheidsdata?

Welke regels of richtlijnen zijn er op het gebied integriteit met betrekking tot publieke overheidsdata?

Dat is redelijk goed. We zijn verplicht aan de wet, dus data wordt binnen deze wettelijke kader gepubliceerd. Er moet worden vermeld waarom iets niet openbaar is, dat kan te maken hebben met privacy, maar er zijn vele redenen, zoals gevoelige informatie van ministeries of de kroon. Het is niet altijd duidelijk wat wel of niet openbaar gemaakt mag worden, bij twijfel wordt het vaak niet openbaar gemaakt, waarbij de beslissing en onderbouwing vaak ontbreken. Als we een open overheid willen, met open data, dan moeten we ook helder aangeven wat wel en niet openbaar is. De data op data.overheid.nl, zoals de data van het Koninklijk Nederlands Meteorologisch Instituut (KNMI), mag voor commerciële doeleinden gebruikt worden. Er is onderzoek gedaan binnen BZK dat aantoont dat er economische groei ontstaat door meer data te delen.

KOOP is bezig met het beschermen van onze open source software. Er zijn markt partijen die er commercieel voordeel uit halen, maar dit is niet de bedoeling. Om dit te voorkomen, werken we met de European Union Public Licence (EUPL), zodat deze sources hergebruikt kunnen worden, maar niet voor commerciële doeleinden.

Compatibility (EA-BD-6)

Welke regels of richtlijnen zijn er op het gebied interoperabiliteit met betrekking tot publieke overheidsdata?

Er zijn standards, zoals de DCAT-standaard van de World Wide Web Consortium (W3C), die de uitwisseling van metadata mogelijk maken. Het Nederlandse toepassingsprofiel 2.0 zal volgens verwachting in de zomer van kracht worden en is ook een Europese standaard. Dit is van belang voor de verzameling van vrijwillige data, omdat de frequentie waarmee de data wordt bijgewerkt niet altijd gelijk is en de formaten soms verschillen. Regels zijn nodig om dit op een efficiënte manier aan te pakken. Er zijn nu verbeteringstrajecten gaande om de data te kunnen samenvoegen.

Algemeen Enterprise Architectuur en Big Data

Algemeen over EA-BD-1 t/m EA-BD-6

KOOP heeft zelf geen BD en er wordt op data.overheid.nl ook geen onderscheid gemaakt tussen data en BD. Voor data.overheid.nl maken we gebruik van de DCAT standaard in wel 90% van de gevallen en richt KOOP zich specifiek op het Europese toepassingsprofiel van DCAT. Het algemene architectuurprincipe is dat, als er een bruikbare standaard beschikbaar is, deze dan gebruikt moet worden, zoals DCAT. Als dit niet het geval is, dan moet er gezocht worden naar een bestaande oplossing, bij voorkeur een internationale of Europese standaard. Als dit niet lukt, dan kan er iets verzonnen worden, maar hierbij is een belangrijk om zoveel mogelijk FAIR en de DCAT richtlijnen te volgen.

D.1.3. Data Governance en Big Data

De interviewvragen over Data Governance (DG) en BD zijn beantwoord door de volgende respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker een
- Beleidsmedewerker twee
- Datamanager

Visie (DG-BD-1)

Wat is de visie met betrekking tot het hergebruik van publieke overheidsdata?

De visie is om de data slechts één keer te verzamelen en zo veel mogelijk te gebruiken, binnen de kaders van privacy en haar beperkingen. Als voorbeeld wordt de interbestuurlijke data strategie genoemd, die het complexe deel van het huidige systeem op een andere manier inricht. Het uitgangspunt voor het bouwen van een federatief data stelsel is dat de data bij de bron bewaard wordt en centraal ontsloten wordt. Als voorbeeld hiervan worden de tien basisregistraties van Nederland genoemd die hun datasets publiceren. Dit is uniek en gebeurt niet in andere landen.

We bekijken eerst in de breedte van de basisregistratie en welke andere datasets daaronder gebracht kunnen worden. Dit kunnen er tientallen tot zelfs honderd zijn. Daarnaast kijken we in diepte alle datasets bij de overheid of zij deze centrale functies ook bij KOOP willen hebben. Als voorbeeld: een centraal advies over gevoelig data waarbij advocaten, juristen en consultants overheidsorganisaties kunnen helpen bij hun dataprojecten. Ook wordt er gekeken naar een catalogusfunctie voor niet alleen voor open data, maar ook gesloten data van alle organisaties, waardoor deze data vindbaar en doorzoekbaar is. KOOP kijken of de principes FAIR - ontsluitbaarheid en interoperabiliteit - hierop aansluiten.

Leiderschap (DG-BD-2)

Welke vormen van leiderschap hebben betrekking tot het hergebruik van publieke overheidsdata?

Er zou meer centrale sturing kunnen zijn, waarbij rekening wordt gehouden met het onderwerp en de domein. Het is op dit moment te abstract en het is begrijpelijker als het gaat over het delen van milieu-, verkeers- en natuurdata. Op dit moment werkt slechts een kleine groep die met open data, het zou meer hapklaar moeten worden aangeboden. Het bestuur in Nederland is extreem versnipperd en er is nu een wens van alle partijen om een centrale regie rol te beleggen bij BZK. Leiderschap wordt steeds belangrijker, maar het open data landschap is decentraal. Het stimuleren van open data ligt meer bij de Open State Foundation (OSF), met als voorlopers het Centraal Bureau voor de Statistiek (CBS) en het Kadaster. Het is in Nederland nog steeds gedecentraliseerd en leiderschap hoeft dan niet hiërarchisch op de hoogste niveau bij het ministerie te zitten.

Financiering Open data (DG-BD-3)

Op welke manier wordt de financiering van publieke overheidsdata geregeld?

De financiering van data.overheid.nl wordt begroot vanuit BZK, hiervoor is er budget beschikbaar onder het data-beleid en onder het openbaar overheidsbeleid. Decentrale overheden en departementen stellen data beschikbaar en financieren dit ook zelf. Het open data-beleid wordt beheerd door BZK als het gaat over de Woo en de technische uitvoering daarvan.

Zelf-ontwikkeld vermogen (DG-BD-4)

Op welke manier wordt het zelf-ontwikkeld vermogen van het hergebruik van publieke overheidsdata vormgegeven?

KOOP bemiddeld tussen vragen om informatie en het aanbod van informatie. Bij de Datamanager van KOOP legt men de informatiebehoefte neer, die vervolgens gaan kijken welke databronnen te vinden zijn en beschikbaar gesteld kunnen worden. Centraal kunnen hergebruik verzoeken worden gedaan via data.overheid.nl, wat een wettelijke verplichting is en waarvoor de verantwoordelijkheid bij de betreffende organisatie ligt.

Communicatie (DG-BD-5)

Op welke manier wordt communicatie rond het hergebruik van publieke overheidsdata vormgegeven?

De belangrijkste communicatie verloopt via data.overheid.nl, evenals Digitale Overheid en datacommunities.nl. KOOP vervult een bemiddelingsrol tussen aanbieders en afnemers, voornamelijk uitgevoerd door de Datamanager. Het CBS beschikt ook over veel data, zoals statistische data.

Interoperabiliteit (DG-BD-6)

Welke data principes worden er toegepast om het hergebruik van publieke overheidsdata te verbeteren?

Voor alle informatie die beschikbaar is op open overheid, moeten duidelijk zijn wat de bruikbaarheid ervan is. De FAIR Data Principles zouden dit makkelijk kunnen maken, daarnaast moeten de licentievoorwaarden voor open data helder zijn. De wettelijke maatregelen van de Woo zal een zware maatregel zijn en gaat naar verwachting veel nieuwe data opleveren. De Organisation for Economic Co-operation and Development (OECD)-data principes opgesteld onder coördinatie van BZK en zijn iets breder dan open data, dus in feite zijn dit de data principes van de overheid. Door gebruik te maken van de DCAT standaarden, spreken alle partijen dezelfde taal. FAIR is belangrijk om goede data te hebben.

Metadata (DG-BD-7)

Welke rol speelt metadata een rol bij het hergebruik van publieke overheidsdata?

Metadata kan worden beschouwd als het middelpunt van de datawereld, omdat het het gemak van het aanleveren en afleveren van data mogelijk maakt. Hoe beter en uniformer de metadata is, des te eenvoudiger is het om data te delen. Een federatief datastelsel en de datacatalogus kunnen hierbij ondersteunen. Binnen KOOP wordt momenteel gewerkt aan het herschrijven van het toepassingsprofiel van DCAT, waaronder voor metadata. Goede metadata is essentieel, omdat je niet elke dataset kunt openen om te controleren of deze aansluit bij je vraag en behoefte.

D.1.4. FAIR Data Management en Big Data

De interviewvragen over FAIR Data Management (FADAM) en BD zijn beantwoord door de volgende respondenten:

- Adviseur Content Standaarden
- Beleidsmedewerker een
- Beleidsmedewerker twee
- Datamanager
- Product Owner DONL

Management Commitment (FADAM-BD-1)

Op welke wijze speelt management commitment een rol bij het hergebruik van publieke overheidsdata?

Het management is zich vaak niet bewust van het feit dat ze in de eerste plaats verantwoordelijk zijn voor het delen van data. Hier is nog heel veel stimulering nodig, waarbij de Woo een belangrijke aanjager is. Als het juridisch van de wet wordt vereist, dan begrijpt een manager dit vaak wel, omdat we ons aan de wet moeten houden. Het management moet vaak de opdracht geven om data openbaar te maken, wat bijdraagt aan de herbruikbaarheid en effectiviteit van data.

Een andere uitdaging voor het management is dat het beschikbaar stellen van data kosten met zich meebrengt zonder iets op te leveren. Hiervoor zou central een budget beschikbaar gesteld kunnen worden. Het commitment van het management is belangrijk, want als zij het niet belangrijk vinden, zal de werkvoet het niet oppakken en omgekeerd, deze combinatie is cruciaal. De overheid beschikt over enorme hoeveelheden data, waaronder gemeenten en departementen, een deel hiervan kan beschikbaar gesteld worden als open data. Later kan er gekeken worden naar het beschikbaar stellen van gesloten data.

Managementplanning (FADAM-BD-2)

Op welke wijze speelt managementplanning een rol bij het hergebruik van publieke overheidsdata?

Het management van KOOP heeft hier vaak onvoldoende aandacht voor de planning van verbetering van de kwaliteit van de data. Binnen BZK streven we ernaar om vijf jaar vooruit een planning te maken, zoals het opzetten van flexibele ontwikkelteams die snelle veranderingen kunnen opvangen en het inrichten van een federatief stelsel. Dit is cruciaal, want *garbage in, garbage out*: als de data van onvoldoende kwaliteit is, verliest het zijn zeggenschap. Hierdoor kunnen verkeerde conclusies worden getrokken uit open data, wat gevaarlijk kan zijn. Management commitment is belangrijker dan de managementplanning.

Vindbaarheid (FADAM-BD-3)

Op welke wijze wordt de vindbaarheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?

Er is behoorlijk geïnvesteerd in de vindbaarheid en daar scoort Nederland wel hoog in. Vindbaarheid is natuurlijk ook een belangrijk onderdeel van data.overheid.nl. Met de stelselcatalogus zijn basisregistraties goed te vinden, hieraan is veel ontwikkeld volgens de twaalf eisen van de wet op de basisregistraties. Het is belangrijk om te benadrukken dat een goed register leidend is voor het uitgeven van unieke identifiers en alles heeft een code of Uniform Resource Identifier (URI).

Er kan een belangrijk onderscheid worden gemaakt tussen enerzijds de omschrijving van het thema van de dataset, het doel en waar het over gaat, en anderzijds de technische aspecten, zoals het bestandsformaat en hoe vaak het wordt bijgewerkt. Het kan voorkomen dat een dezelfde informatie over lantarenpalen bij de ene provincie onder "straatverlichting" valt in de dataset en bij de andere onder "straatmeubilair". Daarom is het belangrijk om goede vindbare metadata te hebben.

Toegankelijkheid (FADAM-BD-4)

Op welke wijze wordt de toegankelijkheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?

Er wordt constant gewerkt aan het beschrijven van de metadata en het beschikbaar te stellen van API's om deze toegankelijk te maken voor machines. De standaarden die hierbij worden gebruikt zijn W3C en Resource Description Framework (RDF). Via de DCAT standaard worden de datasets binnen data.overheid.nl beschreven, inclusief het bestandsformaat. Er zijn aanpassingen nodig voor bepaalde begrippen, omdat het niet altijd duidelijk is wat er in bepaalde bestandsformaten staat. Soms is een speciaal programma, zoals voor geo-data, nodig om deze data te kunnen lezen. Er is nog een grote slag te maken om de herbruikbaarheid van deze data te verbeteren. Vaak kunnen experts deze data wel gebruiken, maar om deze data voor een breder publiek toegankelijk te maken, is het belangrijk om de data op verschillende manieren beschikbaar te stellen.

Interoperabiliteit (FADAM-BD-5)

Op welke wijze wordt de interoperabiliteit van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?

Interoperabiliteit gaat over toegankelijkheid en de mogelijkheid om datasets te combineren. Hiervoor is aandacht nodig en het is belangrijk om te standaardiseren en samen te bespreken welke standaarden bruikbaar zijn voor welke domeinen. Organisaties moeten elkaar vinden en samenkomen, bijvoorbeeld via forum voor standaardisatie. Dit forum beschikt over lijsten met standaarden, hoewel het misschien wenselijk is om meer centrale coördinatie te hebben. Er zijn veel verschillende standaarden, waardoor keuzes moeten worden gemaakt. Er is nog veel verbetering mogelijk, omdat er geen wettelijke basis voor interoperabiliteit is.

Herbruikbaarheid (FADAM-BD-6)

Op welke wijze wordt de herbruikbaarheid van (Meta) data vormgegeven om daarmee publieke overheidsdata te verbeteren?

Als de vindbaarheid verzekerd is, is het herbruikbaar voor een beperkte groep. Maar als de toegankelijkheid en interoperabiliteit verbeteren, zal de herbruikbaarheid alleen maar meer toenemen. Door data open te stellen, zorgen we ervoor dat kan vindbaar is en herbruikbaar door standaardisatie van dataformaten, zoals RDF, JavaScript Object Notation (JSON) en *Terse RDF Triple Language (Turtle)*. De volgende stap is om dingen semantisch interoperabel te maken door gebruik te maken van semantische standaarden met de begripsdefinities die terugkomen in de datacatalogi. Het is belangrijk om te zorgen dat niet elke applicatie een eigen standaard of profiel heeft en dat er gebruik wordt gemaakt van bestaande standaarden. We staan nog aan het begin van dit proces en omdat er veel data wordt ontsloten, heeft dit de hoogste prioriteit. Er zijn verschillende standaarden voor het delen van metadata, de overheid heeft gekozen voor de Europese DCAT-standaard. Hiermee kan alles wat op data.overheid.nl staat een-op-een worden overgezet naar de Europese open data site.

D.1.5. Community Models en Big Data

De interviewvragen over FADAM en BD zijn beantwoord door de volgende respondenten:

- Adviseur Content Standaarden
- Community manager
- Product Owner DONL

Samenwerking (CM-BD-1)

Hoe verloopt de samenwerking met de verschillende datacommunities?

KOOP is in gesprek met veel partijen en er is een wens om elkaar op te zoeken. Er zijn vijf datacommunities opgezet, zoals de datacommunity energie en de datacommunity onderwijs. Waarvan de datacommunity energie het meest actief is, hierin zit onder andere Tennet en energieleveranciers, die om commerciële overwegingen moeite hebben om data te delen. Zij hebben ook grote uitdagingen zoals de CO2 uitdaging en de transformatie en vinden het lastig om data te ontsluiten, maar hebben wel toegevoegde waarde.

De datacommunities hebben geen contact met elkaar, alleen als personen in meerdere datacommunities zitten. Als iemand een specialistisch onderwerp heeft, kan men een community starten en mensen kunnen vrijblijvend lid worden. Daar kunnen vragen gesteld en informatie worden gedeeld, net zoals op LinkedIn en Facebook groepen. Communities zijn met name ontstaan tijdens de Corona periode als een *"nice-to-have"*, omdat er geen congressen en bijeenkomsten konden plaatsvinden over bepaalde onderwerpen. Hierdoor is er wel een behoefte ontstaan, maar deze neemt wel af. Het is een verlenging van een congres, waar meer mensen over de onderwerpen kunnen praten, bijvoorbeeld bij een congres over energie komen ongeveer 900 personen.

Het vaststellen van standaarden die voor verschillende gremia moeten gelden is lastig. KOOP heeft een eigen standaardisatieraad en een gemeente ook, waardoor er aan twee standaarden gehouden moet worden. Dit maakt het lastig.

Vaardigheden en training (CM-BD-2)

Wordt er bepaalde vaardigheden verwacht van de verschillende datacommunities en zijn daarvoor trainingen ingericht?

Er is onlangs een nieuwe Community manager aangesteld voor data.overheid.nl. In het afgelopen half jaar zijn er geen trainingen gegeven, maar dat is ook niet verplicht. Belangrijke vaardigheden zijn een open *mindset* een *up-to-date* inhoudelijke kennis, evenals het vermogen om je in te leven in een ander domein of sector. De communities zijn heel laagdrempelig en het enige wat je nodig hebt is om je aan te melden en iets te kunnen posten. Er is een uitleg filmpje beschikbaar over de datacommunities.nl en hoe ze werken en deze is geïntegreerd in data.overheid.nl. Het delen van data is vrij complex. Binnen de datacommunities.nl van data.overheid.nl overwegen we of we een soort *road-show* kunnen organiseren om mensen te ondersteunen. Dit kan door hen stap voor stap te begeleiden, instructies te geven en hen op een laagdrempelige manier aan te moedigen om gebruik te maken van dit systeem. We merken dat veel technische mensen het systeem gebruiken en voor hen is het geen probleem, maar dat geldt niet voor nieuwkomers zoals de gemiddelde beleidsmedewerker bij een gemeente.

Openheid (CM-BD-3)

Welke mate van openheid richting de verschillende datacommunities wordt er nagestreefd?

Vroeger was er veel angst om data te delen, maar nu is er besef dat het moet. De vraag is nu hoe we dit op een zorgvuldige manier kunnen organiseren. Er is al verschuiving gaande "het mag niet naar buiten" naar "het gaat naar buiten, maar hoe". KOOP erkent het belang van de Wet openbaarheid van bestuur (Wob) en de bijbehorende kosten. De Woo probeert een grote stap te zetten in de openheid van data. Hoewel niet alle data helemaal open is, maar alles wat open data is, moet alles wat als open data wordt beschouwd publiekelijk toegankelijk en herbruikbaar zijn en gemakkelijk vindbaar zijn. Het opmerkelijke is dat de datacommunities worden afschermd met een gebruikersnaam en wachtwoord, wat willicht een drempel vormt. Binnen de datacommunities worden zaken ook gedeeld.

Technische infrastructuur (CM-BD-4)

Worden er technische infrastructuur beschikbaar gesteld aan de verschillende datacommunities?

De basis principe is om de data bij de bron te houden, maar er is technische infrastructuur beschikbaar op data.overheid.nl wanneer dit niet mogelijk is. Voor documenten zijn we op zoek naar een manier om deze data beschikbaar te maken, bijvoorbeeld via het Platform Open Overheidsinformatie (PLOOI). De tool "Huddle" wordt gebruikt om de datacommunities te ondersteunen, waarin mensen zelf groepen kunnen aanmaken en discussiëren over onderwerpen zoals zonnepanelen. Data.overheid.nl kan ook worden beschouwd als een tool.

Gemeenschappelijke participatie (CM-BD-5)

Hoe wordt er aandacht besteed aan gemeenschappelijke participatie met de verschillende datacommunities?

Data.overheid.nl houdt regelmatig bijeenkomsten met gebruikers en stakeholders, en de indruk is dat dit zeer op prijs wordt gesteld. Het is erg belangrijk dat de betrokken partijen kunnen uitwisselen van kennis en ervaringen kunnen uitwisselen en zodra de corona achter de rug is, moeten we weer seminars organiseren om elkaar te ontmoeten. Bij datacommunities zijn er ambassadeurs die het persoonlijk belangrijk vinden dat deze community bestaat en zorgen ervoor dat vragen snel beantwoord worden. Een van de doelen is om deze actieve personen samen te brengen om te peilen wat ze vinden het platform vinden en mogelijk verbeteringen verbeteringen aangebracht kunnen worden. Hierdoor kunnen de ambassadeurs van de datacommunities kennis en vaardigheden uitwisselen en problemen oplossen waar anderen mogelijk vastlopen op data.overheid.nl. Gemeenschappelijke participatie komt echter in de praktijk niet veel voor.

Het hoofddoel van data.overheid.nl is het verkrijgen en verstrekken van data die door de overheid is betaald. Rapporten die zijn gemaakt kunnen worden opgehaald van data.overheid.nl en als deze "up-to-data" zijn, hoeft je zelf geen onderzoek meer uit te voeren, als je Toegepast-Natuurwetenschappelijk Onderzoek (TNO) bent.

Economisch en bedrijfsmodel (CM-BD-6)

Welke mate worden de verschillende datacommunities ondersteund op het gebied van economisch en bedrijfsmodellen?

Dit wordt niet uitgevoerd door KOOP. Het wordt ook niet aangetroffen in de praktijk, zoals van de datacommunity energie. De data op data.overheid.nl zijn openbaar, onze stakeholders zijn ook ondernemers die de open data kunnen gebruiken voor economische doeleinden.

Juridische en ethische kwesties (CM-BD-7)

Hoe wordt met juridische en ethische kwesties omgegaan met de verschillende datacommunities?

De AVG moet nauwlettend worden gevolgd, bij documenten is er veel aandacht voor persoonsgegevens die weg gelakt moeten worden. Voor Wob-, Woo-verzoeken worden deze door KOOP gepubliceerd, de aanleverende partij is zeer zorgvuldig met het beschermen van deze persoonsgegevens. Ethische kwesties spelen tot nu toe zelden een rol omdat het voornamelijk gaat om openbare informatie. Het betekent dat er geen aandacht voor is of dat dit genegeerd wordt, het is tot nu toe nog niet aan de orde geweest. Voor de datacommunities.nl zijn er regels opgesteld waaraan moet worden voldaan, bij nietnaleving krijg men een waarschuwing of wordt de toegang ontnomen. Hierbij kan gedacht worden aan ongepaste taalgebruik of discriminatie. De datacommunities.nl zijn nog in de kinderschoenen en als ze drukker worden, worden deze regels belangrijker en zullen ze misschien worden bijgesteld. Binnen een datacommunity energie zijn een aantal juristen die hun uiterste best doet om partijen te bewegen om de uitdagingen waar ze voor staan in kader van de energietransitie en het delen van data mogelijk te maken.

Academische cultuur (CM-BD-8)

Hoe is de academische cultuur vormgegeven bij de verschillende datacommunities?

Het initiatief voor het platform LD komt van de Universiteit van Amsterdam (UvA) en was sterk academisch georiënteerd. Hoewel ze dit nog steeds volgen, is het platform LD nu in handen van een commerciële partij, waardoor de academische drive is verminderd. Alle deelnemers zijn wel hoog opgeleid, wat betekent dat het academische niveau van een oplossing niet in twijfel getrokken kan worden.

Bij de datacommunities is er een beloningsmodel ingebouwd in de tool. Bijvoorbeeld, als je veel post, dan kun je een hoger level en beloningen verdienen. Dit geeft inzicht in de activiteit van iemand op het platform.

Er zijn af en toe verzoeken voor data vanuit de wetenschappelijke wereld, waaronder wetenschappers en hoogleraren, voor nieuwe datasets op data.overheid.nl, maar dit betreft slechts een kleine groep.

Participatie (CM-BD-9)

Hoe wordt de participatie met de verschillende datacommunities vormgegeven?

Er vinden veel *ad hoc* overleggen plaats, van operationele overleggen op lagere niveaus tot weinig overleggen op bestuurlijk niveau. Participatie is het uitwisselen van informatie tussen de ambassadeurs en trekkers van elke datacommunity, om te bekijken wat wel of niet goed gaat in verband met data.overheid.nl en om verbeteringen aan te brengen en veranderingen door te voeren. Daarnaast verzenden de trekkers van een datacommunity nieuwsbrieven naar de verschillende leden.

Engagement (CM-BD-10)

Op welke wijze wordt engagement bewerkstelligd bij de verschillende communities?

Dit wordt voornamelijk gedaan via de datacommunities.nl op data.overheid.nl, ze organiseren de onderwerpen, zoals voor de stelselcatalogi, mede omdat dit is voorgeschreven door de Omgevingswet. Ze proberen ook actuele en interessante onderwerpen op de agenda te zetten, zoals de datacommunity energie voor een bijeenkomst met 150 deelnemers. Er zijn plannen om een soort marketingplan uit te rollen en er wordt een soort *road-show* georganiseerd in samenwerking met de Vereniging Nederlandse Gemeenten (VNG) en BZK kan een congres over data en BD organiseren. Het doel is om mensen kennis te laten maken en te laten communiceren, en het is fijn als er sprekers aanwezig zijn over BD en data.overheid.nl. Ook is er de mogelijkheid om kern inspecteurs bij de gemeenten op te leiden of een korte cursus te geven hoe ze data op data.overheid.nl kunnen zetten, waarmee we misschien aan de slag gaan in het najaar.

Binnen KOOP wordt er gebrainstormd hoe ze de stakeholders willen bereiken en wie baat heeft bij data.overheid.nl en wie verplicht is data te delen, zoals gemeentes. [Data.overheid.nl](http://data.overheid.nl) moet het platform worden voor alle data en het is belangrijk dat er een goede zoekfunctie is, en de datacommunities kunnen daarbij helpen.

Sociale media speelt ook een belangrijke rol en door posts op LinkedIn over datacommunities, kun je meer mensen naar die datacommunity trekken.

Stimulansen (CM-BD-11)

Op welke wijze wordt de verschillende communities gestimuleerd voor data hergebruik?

De datacommunities op data.overheid.nl worden beleidsmatig gestimuleerd. Andere samenwerking ontstaat van onderaf als er een behoefte is aan samenwerking of als men op zoek is naar een contactpersoon bij de andere partij. Dit gebeurt niet vaak, maar er is mogelijk nog ruimte voor verbetering. De datacommunities op data.overheid.nl zijn hiervoor opgericht en worden door ambassadeurs actief gehouden. Dit is een uitdaging, vooral voor de Community manager druk mee bezig is. Veel mensen binnen de datacommunities zijn druk met hun werk en hebben weinig tijd over.

De verwachting is dat de Woo een grote stimulans zal zijn om data openbaar te maken.

Community Governance Models (CM-BD-12)

Worden door de verschillende datacommunities Governance models gebruikt?

Community Governance Models, die bedoeld zijn om de inrichting en volwassenheid van communities te sturen, worden niet door het management KOOP gebruikt. Er wordt nog steeds vaak in de waan van de dag geleefd.

D.2. Resultaten triangulatie antwoorden

Hieronder de controleantwoorden waarin de uitslagen van de vijfpunts-Likertschaal voor elk kernbegrip (EA, EA-DG, FADAM-BD en Community Models (CM)-BD) wordt gepresenteerd voor alle respondenten. Deze uitslagen zijn tijdens de interviews verkregen door de respondenten te vragen, als onderdeel van de triangulatie voor de kwalitatief onderzoeksvragen in paragraaf 'D.1-Samenvatting transcripties', met als doel antwoord te krijgen op de toepasbaarheid van de kernbegrippen.

D.2.1. Analyse kernbegrippen en dimensies

Op basis van de kernbegrippen en Dimensies zoals deze zijn aangegeven in de transcripties van het interview van alle respondenten is het onderstaande tabel 'D.1-Analyse kernbegrippen' ingevuld.

Aan de respondenten is gevraagd in hoeverre een dimensie wordt toegepast in de huidige situatie binnen de case-organisatie, dit op basis van een vijfpunts-Likertschaal.

De huidige situatie matrix is: niet toegepast – nauwelijks toegepast – wordt wel toegepast – is verplicht – is randvoorwaardelijk.

Hierbij is ter controle ook gevraagd aan te geven op een vijfpunts-Likertschaal, de toepasbaarheid van een dimensie om daarmee de herbruikbaarheid van publieke overheidsdata te kunnen verbeteren [Likert, 1932].

De toepasbaarheid matrix is: niet toepasbaar – beperkt toepasbaar – goed toepasbaar – noodzakelijk – randvoorwaardelijk.

De totaal score is een gewogen gemiddelde van alle respondenten met een verdeling van: $-2 - -1 - 0 - +1 - +2$. Dus bij een score van 0, hoort "wordt wel toegepast" voor de huidige situatie binnen de casus-organisatie en "goed toepasbaar" voor de toepasbaarheid van die dimensie om daarmee de herbruikbaarheid van publieke overheidsdata te kunnen verbeteren.

Kernbegrip	Dimensie							Gemiddeld alle	Totaal
		Beleidsmedewerker een	Beleidsmedewerker een	Product Owner DONL	Datamanager	Community manager	Adviseur Content Standaarden		
Toepassen van EA principes en richtlijnen	EA nu	-	0	-1	0	-	0	-0,3	Wordt beperkt toegepast
	EA gewenst	-	+1	+1	+2	-	+1	+1,3	Tussen noodzakelijk en randvoorwaardelijk
	EA-BD nu	-	-	0	-1	-	0	-0,3	Wordt beperkt toegepast
	EA-BD gewenst	-	-	+2	0	-	0	+0,7	Tussen toepasbaar en randvoorwaardelijk
Toepassen van DG principes en richtlijnen	DG-BD nu	0	+0,5	-	-	-	0	+0,2	Wordt toegepast
	DG-BD gewenst	+1	0	-	-	-	0	+0,4	Goed toepasbaar misschien noodzakelijk
Toepassen van FADAM principes en richtlijnen	FADAM-BD nu	-1	0	-1	-1	-	0	-0,6	Nauwelijks toegepast
	FADAM-BD gewenst	+2	0	0	+1	-	+1	+0,8	Noodzakelijk
Ontwikkelen van Best Practices CM	CM-BD nu	-	-	0	-	-1	-1	-0,7	Nauwelijks toegepast
	CM-BD gewenst	-	-	0	-	+2	0	+0,7	Verplicht

Tabel D.1: Analyse kernbegrippen

D.3. Analyse documentatie

Hieronder staan de metadata van de documentatie en een samenvattingen van de mogelijke relevante data. Dit betreft de volgende documenten:

- Wet openbaarheid van bestuur (Wob)
- Wet open overheid (Woo)
- Wet hergebruik van overheidsinformatie (Who)
- Data Catalog Vocabulary (DCAT)

D.3.1. Wet openbaarheid van bestuur

In tabel 'D.2-Analyse document Wet openbaarheid van bestuur' zijn de metadata van de Wob te vinden.

Document	Wet openbaarheid van bestuur (Wob)
Titel	Wet openbaarheid van bestuur (Wob)
Ontvangst informatie	BZK en KOOP
Auteur	BZK
Publicatiedatum	28-07-2018
Status	Geldend van 28-07-2018 t/m 30-04-2022
Versie	28-07-2018
Taal	Nederland
Beschikbaarheid	https://wetten.overheid.nl/BWBR0005252/2018-07-28
Aantal citaten	–
Score document	Relevant, wetgeving
Totaal score document	Meenemen
Gemiddelde score authenticatie	Hoog, wetgeving
Gemiddelde score datum publicatie	–
Gemiddelde score relevantie	Relevant
Gemiddelde score scope	Hoog
Relatie document met de onderzoeksvragen	Was de wetgeving waarin de openbaarheid van overheidsinformatie in was geregeld.
Hoe wordt het document gebruikt door de actoren	Als wetgeving, basis voor het openbaar maken van overheidsinformatie.
Samenvatting:	
Openbaarheid: <ul style="list-style-type: none"> • Een bestuursorgaan verstrekt bij de uitvoering van zijn taak, onverminderd het elders bij wet bepaalde, informatie overeenkomstig deze wet en gaat daarbij uit van het algemeen belang van openbaarheid van informatie. • Het bestuursorgaan draagt er zo veel mogelijk zorg voor dat de informatie die het overeenkomstig deze wet verstrekt, actueel, nauwkeurig en vergelijkbaar is. 	
Vervangen door de Woo, zie tabel 'D.3-Analyse document Wet open overheid'.	

Tabel D.2: Analyse document Wet openbaarheid van bestuur

D.3.2. Wet open overheid

In tabel 'D.3-Analyse document Wet open overheid' zijn de metadata van de Woo te vinden.

Document	Wet open overheid (Woo)
Titel	Wet open overheid (Woo)
Ontvangst informatie	BZK en KOOP
Auteur	BZK
Publicatiedatum	01-08-2022
Status	Geldig van 01-08-2022 t/m heden
Versie	01-08-2022
Taal	Nederland
Beschikbaarheid	https://wetten.overheid.nl/BWBR0045754/2022-08-01
Aantal citaten	-
Score document	Relevant, wetgeving
Totaal score document	Meenemen
Gemiddelde score authenticatie	Hoog, wetgeving
Gemiddelde score datum publicatie	-
Gemiddelde score relevantie	Relevant
Gemiddelde score scope	Hoog
Relatie document met de onderzoeksvragen	Huidige wetgeving waarin de openbaarheid van overheidsinformatie in was geregeld.
Hoe wordt het document gebruikt door de actoren	Als nieuwe wetgeving, basis voor het openbaar maken van overheidsinformatie.
Samenvatting:	
<p>Openbaarmaking:</p> <p>Rechten, recht op toegang.</p> <p>Eenieder heeft recht op toegang tot publieke informatie zonder daartoe een belang te hoeven stellen, behoudens bij deze wet gestelde beperkingen.</p> <p>Openbaarmaking uit eigen beweging.</p> <ol style="list-style-type: none"> 1. Het bestuursorgaan dat het rechtstreeks aangaat, maakt bij de uitvoering van zijn taak uit eigen beweging de bij het bestuursorgaan berustende informatie neergelegd in documenten voor eenieder openbaar, indien dit zonder onevenredige inspanning of kosten redelijkerwijs mogelijk is, behoudens voor zover de artikelen 5.1, eerste, tweede en vijfde lid, en 5.2 aan openbaarmaking in de weg staan of met de openbaarmaking geen redelijk belang wordt gediend. Deze informatie betreft in ieder geval informatie over het beleid, inclusief de voorbereiding, uitvoering, naleving, handhaving en evaluatie. 2. Het bestuursorgaan doet bij een gedeeltelijke niet-openbaarmaking hiervan mededeling gelijktijdig met de openbaarmaking. 3. Documenten als bedoeld in het eerste lid worden niet openbaar gemaakt dan nadat belanghebbenden die naar verwachting bedenkingen zullen hebben tegen openbaarmaking, in de gelegenheid zijn gesteld binnen een door het bestuursorgaan gestelde termijn hun zienswijze naar voren te brengen. 4. Het bestuursorgaan deelt een belanghebbende mede dat toepassing wordt gegeven aan het eerste lid, onder vermelding van het tijdstip van openbaarmaking en de openbaar te maken documenten. De mededeling wordt gelijkgesteld met een besluit. <p>Openbaarmaking op verzoek.</p> <ol style="list-style-type: none"> 1. Eenieder kan een verzoek om publieke informatie richten tot een bestuursorgaan of een onder verantwoordelijkheid van een bestuursorgaan werkzame instelling, dienst of bedrijf. In het laatste geval beslist het verantwoordelijke bestuursorgaan op het verzoek. 2. Een verzoek kan mondeling of schriftelijk worden ingediend en kan elektronisch worden verzonden op de door het bestuursorgaan aangegeven wijze. 3. De verzoeker behoeft bij zijn verzoek geen belang te stellen. 4. De verzoeker vermeldt bij zijn verzoek de aangelegenheid of het daarop betrekking hebbende document, waarover hij informatie wenst te ontvangen. 5. Indien een verzoek te algemeen geformuleerd is, verzoekt het bestuursorgaan binnen twee weken na ontvangst van het verzoek de verzoeker om het verzoek te preciseren en is het de verzoeker daarbij behulpzaam. 6. Het bestuursorgaan kan besluiten een verzoek niet te behandelen, indien de verzoeker niet meewerkt aan een verzoek tot precisering als bedoeld het vijfde lid. In afwijking van artikel 4:5, vierde lid, van de Algemene wet bestuursrecht wordt het besluit om het verzoek niet te behandelen aan de verzoeker bekendgemaakt binnen twee weken nadat het verzoek is gepreciseerd of nadat de daarvoor gestelde termijn ongebruikt is verstreken. 7. Een verzoek om informatie wordt ingewilligd met inachtneming van het bepaalde in hoofdstuk 5. 	

Tabel D.3: Analyse document Wet open overheid

D.3.3. Wet hergebruik van overheidsinformatie

In tabel 'D.4-Analyse document Wet hergebruik van overheidsinformatie' zijn de metadata van de Who te vinden.

Document	Wet hergebruik van overheidsinformatie (Hwo)
Titel	Wet hergebruik van overheidsinformatie (Hwo)
Ontvangst informatie	BZK en KOOP
Auteur	BZK
Publicatiedatum	01-07-2021
Status	Geldend van 01-07-2021 t/m heden
Versie	01-07-2021
Taal	Nederland
Beschikbaarheid	https://wetten.overheid.nl/BWBR0036795/2021-07-01
Aantal citaten	-
Score document	Relevant, wetgeving
Totaal score document	Meenemen
Gemiddelde score authenticatie	Hoog, wetgeving
Gemiddelde score datum publicatie	-
Gemiddelde score relevantie	Relevant
Gemiddelde score scope	Hoog
Relatie document met de onderzoeksvragen	Huidige wetgeving waarin de openbaarheid van overheidsinformatie in was geregeld.
Hoe wordt het document gebruikt door de actoren	Als nieuwe wetgeving, basis voor het openbaar maken van overheidsinformatie.
Samenvatting:	
<p>Aanhef:</p> <p>Alzo Wij in overweging genomen hebben, dat het noodzakelijk is regels te stellen over het hergebruik van overheidsinformatie in verband met de implementatie van Externe link: Richtlijn nr2013/37/EU van het Europees Parlement en de Raad van 26 juni 2013 tot wijziging van Externe link: Richtlijn 2003/98/EG inzake het hergebruik van overheidsinformatie (PbEU 2013, L 175/1);</p> <p>Hergebruik:</p> <p>hergebruik: het gebruik van informatie, neergelegd in documenten berustend bij een met een publieke taak belaste instelling, voor andere doeleinden dan het oorspronkelijke doel binnen de publieke taak waarvoor de informatie is geproduceerd, anders dan de uitwisseling van informatie tussen met een publieke taak belaste instellingen onderling uitsluitend met het oog op de vervulling van hun publieke taken;</p>	

Tabel D.4: Analyse document Wet hergebruik van overheidsinformatie

D.3.4. Data Catalog Vocabulary

In tabel 'D.5-Analyse document Data Catalog Vocabulary' zijn de metadata van de DCAT te vinden.

Document	Data Catalog Vocabulary (DCAT)
Titel	Data Catalog Vocabulary (DCAT)
Ontvangst informatie	BZK en KOOP
Auteur	W3C
Publicatiedatum	04-02-2020
Status	Geldig
Versie	04-02-2020
Taal	Engels
Beschikbaarheid	https://www.w3.org/TR/vocab-dcat-2/
Aantal citaten	–
Score document	Relevant, internationale en nationale standaard
Totaal score document	Meenemen
Gemiddelde score authenticatie	Hoog, standaard
Gemiddelde score datum publicatie	–
Gemiddelde score relevantie	Relevant
Gemiddelde score scope	Hoog
Relatie document met de onderzoeksvragen	Internationale en nationale standaard.
Hoe wordt het document gebruikt door de actoren	Als nationale standaard, basis voor het kunnen delen van overheidsinformatie.
Samenvatting:	
<p>Abstract:</p> <p>DCAT is an RDF vocabulary designed to facilitate interoperability between data catalogs published on the Web. This document defines the schema and provides examples for its use. DCAT enables a publisher to describe datasets and data services in a catalog using a standard model and vocabulary that facilitates the consumption and aggregation of metadata from multiple catalogs. This can increase the discoverability of datasets and data services. It also makes it possible to have a decentralized approach to publishing data catalogs and makes federated search for datasets across catalogs in multiple sites possible using the same query mechanism and structure. Aggregated DCAT metadata can serve as a manifest file as part of the digital preservation process.</p>	

Tabel D.5: Analyse document Data Catalog Vocabulary