



Universiteit
Leiden
The Netherlands

**Advisering op het Gebied van AI, een Pad naar Inzicht en Innovatie:
Een verkennend onderzoek naar de mogelijkheden voor een
adviserend orgaan op het gebied van AI**

Nagtegaal, Luuk

Citation

Nagtegaal, L. (2024). *Advisering op het Gebied van AI, een Pad naar Inzicht en Innovatie: Een verkennend onderzoek naar de mogelijkheden voor een adviserend orgaan op het gebied van AI.*

Version: Not Applicable (or Unknown)

License: [License to inclusion and publication of a Bachelor or Master Thesis, 2023](#)

Downloaded from: <https://hdl.handle.net/1887/4149987>

Note: To cite this publication please use the final published version (if applicable).

Advisering op het Gebied van AI, een Pad naar Inzicht en Innovatie

Een verkennend onderzoek naar de mogelijkheden voor een adviserend orgaan op het gebied van AI



Universiteit Leiden Campus Den Haag

Naam student: Luuk Thomas Nagtegaal
Studentnummer: s3972410
Faculteit: Governance and Global Affairs
Studie (MSc.): Management van de Publieke Sector - Politiek, Beleid en Management
Begeleider: Prof.dr.ing. A.J. Klievink

IJsselstein, 31 mei 2024

Deze masterthesis voor de studie Management van de Publieke Sector is geschreven onder de verantwoordelijkheid van de Universiteit Leiden. Het copyright berust bij de auteur. Zowel de Universiteit Leiden als de auteur verklaren, dat zij eventuele gegevens van derden die voor deze afstudeerscriptie zijn gebruikt en die door deze derden als vertrouwelijk zijn aangemerkt, als zodanig zullen behandelen.

Samenvatting

Het is vrijdag 15 januari 2021 als minister-president Mark Rutte in een persconferentie aankondigt dat hij het ontslag van het voltallige kabinet heeft aangeboden aan de koning. De oorzaak van het vallen van kabinet Rutte III is het rapport *Ongekend onrecht*, dat een maand eerder werd gepubliceerd door de Parlementaire ondervragingscommissie Kinderopvangtoeslag (NOS, 2021a). In haar rapport is de ondervragingscommissie spijkerhard, zo stelt zij dat er grondbeginselen van de rechtsstaat zijn geschonden en dat alle drie de machten zich dat aan mogen rekenen (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020). De vragen: *hoe heeft het zo ver kunnen komen?* en *hoe kon het zo mis gaan?* domineren weken, maanden en zelfs jaren later nog steeds in de media (oa RTL, 2020. NOS, 2021b. NOS, 2022).

Het toeslagenschandaal laat zien hoe pijnlijk het gebruik van AI fout kan gaan, maar dat doemscenario gaat gelukkig niet op voor alle AI-toepassingen die gebruikt worden. Het gebruik van AI-toepassingen brengt veel kansen en mogelijkheden met zich mee, in de positieve zin van het woord. Zo kan het ingezet worden binnen de gezondheidszorg en maakt het het doen van inkomstenbelasting een stuk makkelijker (Henman, 2020, pp. 210-213). Tevens is er in de academische discussie over hoe een overheid om dient te gaan met AI veel ruimte voor zowel de voor- als nadelen van dit AI-gebruik. Daar worden dan vaak governance modellen of frameworks aan gekoppeld (oa. Almeida & Gasser, 2017; Sturm, Weyer & Wirtz, 2020). De behoefte aan dit soort modellen wordt getekend door de hoeveelheid wet- en regelgeving die men tracht te creëren voor AI, bijvoorbeeld de Europese AI-act. Die behoefte naar institutionele sturing op AI blijft groeien, zo stuurde Von der Leyen vorig jaar een brief waarin zij een Europese AI-raad aankondigde. Ook in Nederland is deze behoefte groeiende, zo werd de motie van de leden Dekker-Abdulaziz en Rajkowski die het kabinet oproept om een Nederlandse AI-raad te onderzoeken aangenomen. Daarmee groeit ook de behoefte naar (academisch) onderzoek op het snijvlak van advisering en AI. Feit is echter dat de literatuur over hoe overheden zich kunnen laten adviseren over het gebruik van AI nauwelijks bestaat. Deze thesis beoogt een begin te zijn om die mismatch te ondervangen, en daarmee het snijvlak van advisering en AI (hoger) op de academische agenda te zetten.

Er is academische literatuur over advisering aan de overheid in allerlei vormen, van factoren die meespelen bij het inrichten van een dergelijk orgaan tot manieren van kennismanagement (oa. Acosta et al, 2022; Osborne, 2004). Die literatuur gaat zelfs een stap verder, naar beleidsbeïnvloeding (bijv. Aarts et al, 2022). Echter ontbreekt een academische blauwdruk wanneer men een adviesorgaan wil inrichten; er zijn weinig tot geen peer-reviewed artikelen over wat nou wel en niet effectief is op dit vlak. Daarnaast is er literatuur waarin een (academisch) fundament wordt gelegd voor wet- en regelgeving van AI. Die literatuur varieert van de vraag of er überhaupt aparte wet- en regelgeving voor AI moet zijn (Chauhan, 2022) tot complete denkrichtingen en uitgewerkte governance modellen voor deze wet- en regelgeving (oa. Almeida & Gasser, 2017; Choi & Park, 2023). Dit onderzoek beoogt deze literatuur te koppelen, zodat er een begin wordt gemaakt om dit kennisprobleem op te lossen.

Deze thesis is daarmee een verkenning naar de mogelijkheden van een adviesorgaan op het gebied van AI.

De hoofdvraag van dit onderzoek is: *Wat zijn de mogelijkheden voor het inrichten van een adviserend orgaan op het gebied van AI?*

Om die hoofdvraag te beantwoorden is eerst een literatuurstudie gedaan. Daarin zullen de hierboven beschreven onderwerpen aan bod komen, zodat er een overzicht is over de verschillende facetten uit de literatuur waar dit onderzoek mee te maken heeft. Uit dit overzicht zal een topiclijst worden gedestilleerd om de verzamelde data mee te analyseren. Deze data wordt verzameld middels interviews en documentenanalyses, en is daarmee kwalitatief. De geanalyseerde resultaten zullen gebruikt worden om de empirische verwachtingen te ontcrachten of bevestigen, waarna in de conclusie de hoofdvraag beantwoord wordt.

Uit de resultaten is gebleken dat een tweetal mogelijkheden het meest naar boven komen, en daarmee wellicht het meest voor de hand liggen. Dat zijn het onderbrengen van het thema AI bij een bestaand adviesorgaan en het creëren van een coördinerend orgaan op het gebied van advisering over AI. De onderbouwing voor de tweede mogelijkheid is meer het zwaartepunt dan die van de eerste, en komt daardoor ook veel meer terug in de resultaten. De argumentatie hiervoor zit hem in de mogelijkheid om enerzijds zelf op een hoger abstractieniveau te adviseren en anderzijds meer specifieke vragen door te zetten naar organen met expertise op het gevraagde advies. Kortom, een makelaarsfunctie dus.

Deze conclusie kan echter niet heel hard getrokken worden, gezien het geringe aantal respondenten en de brede insteek van dit onderzoek. Daarom zijn de aanbevelingen voor vervolgonderzoek om de twee genoemde mogelijkheden verder te onderzoeken alsmede te onderzoeken wat nou echt effectieve vormen van expertadviesing aan de overheid zijn.

Voorwoord

Beste lezer,

Voor u ligt mijn masterthesis ter afronding van de master *Management van de Publieke Sector - Politiek, Beleid en Management*, dit zal het laatste zijn wat u leest van de student Luuk Nagtegaal. Deze thesis vormt namelijk niet alleen het slotstuk van mijn tijd aan de Universiteit Leiden, maar tevens dat van mijn studententijd. De eerlijkheid gebiedt te zeggen dat ik het studeren zal missen. Hoewel het schrijven van papers of het maken van tentamens misschien niet mijn favoriete bezigheden waren, heb ik genoten van de eindeloze interessante gesprekken en discussies met mijn studiegenoten en hoogleraren. Met het einde van mijn studententijd begint ook een nieuwe tijd voor mij, die van het werkzame leven. Ik kijk er ontzettend naar uit om mijn opgedane kennis en kunde in te kunnen zetten voor de Nederlandse samenleving, de samenleving die voor mij het studeren hoofdzakelijk heeft mogelijk gemaakt. De afgelopen vijf jaar heb ik kennis mogen opdoen aan de Universiteit Utrecht en de Universiteit Leiden. Een studententijd die gekenmerkt is door de corona-pandemie, maar die mij desondanks bij zal blijven als misschien wel de leukste periode van mijn leven. Vijf jaren gevuld met vakken, talloze colleges en werkgroepen, stages en mooie steden met ieder een eigen dynamiek. Ik ben daar ontzettend dankbaar voor.

Dankbaar ben ik ook voor de begeleiding die ik tijdens het schrijven van deze thesis heb mogen krijgen van prof.dr.ing. A.J. Klievink. Zijn manier van feedback geven, welke te omschrijven is als gericht, constructief en daarmee treffend, maakte het proces voor mij ontzettend leerzaam. Als bepaalde delen wat minder voorspoedig verliepen, was hij de rots in de branding waar ik op kon bouwen. Bram, heel erg bedankt daarvoor.

Tot slot ben ik dankbaar voor de kansen die ik heb mogen krijgen. Toen ik als twaalfjarig jongetje mijn eerste stap in mijn middelbare school zette, durfde ik niet te dromen van de plek waar ik nu sta. Ik durfde niet te dromen van de mogelijkheid om aan de universiteit te studeren. Elf jaar later kan ik als eerste generatie student van mijn familie met trots zeggen dat het gelukt is. Het is ook die familie die ik dankbaar ben voor het motiveren van mij, in voor- en tegenspoed. Het was mijn opa die mij op het hart drukte dat het zou lukken, dat ik het allemaal zou halen en dat ik hem trots zou maken ongeacht het resultaat. Het waren mijn ouders die mij altijd gestimuleerd hebben om alles eruit te halen wat erin zat. De afgelopen jaren ben ik mij steeds meer bewust geworden van het feit dat ik zonder hen, zonder hun steun, zonder het fijne thuis wat ik had, dit niet had kunnen doen. Zonder de onvoorwaardelijke liefde, met een vleugje engelengeduld, die ik van mijn ouders heb mogen krijgen, had ik niet gestaan op de plek waar ik vandaag de dag sta. Het waren de motiverende woorden van mijn opa en mijn ouders die mij in staat stelden dit te doen. Ik draag deze thesis daarom op aan hen. Opa, papa, mama, dit is voor jullie.

Luuk Nagtegaal
IJsselstein, 31 mei 2024

Inhoudsopgave

Samenvatting	1
Voorwoord	3
Inhoudsopgave	4
1. Inleiding	6
1.1 Aanleiding	6
1.2 Probleemstelling (vraag- en doelstelling)	7
1.3 Maatschappelijke relevantie	8
1.4 Wetenschappelijke relevantie	8
1.5 Leeswijzer	8
2. Theoretisch kader	9
2.1 Inleiding	9
2.2 AI	10
2.3 Regulering en Governance van AI	10
2.3.1 Waarom regulering?	11
2.3.2 Governancemodellen	12
2.4 Adviesorganen	14
2.4.1 Theoretische beschrijvingen	14
2.4.2 Kennismanagement	15
2.5 Theoretische lens & empirische verwachtingen	16
3. Methoden & Technieken	18
3.1 Inleiding	18
3.2 Onderzoeksstrategie	18
3.3 Dataverzameling en data-analyse	19
3.3.1 Semi-gestructureerde interviews	19
3.3.2 Documentanalyse	20
3.3.3 Data-analyse	20
3.4 Operationalisatie theoretische concepten	21
3.5 Validiteit en betrouwbaarheid	21
3.5.1 Validiteit	21
3.5.2 Betrouwbaarheid	22
4. Resultaten	23
4.1 Governance van AI	23
4.2 Functioneren en effectiviteit van adviesorganen	25
4.3 Advisering op het gebied van AI	27
4.3.1 Inzichten uit de data	27
4.3.2 Bestaande manieren van advisering op het gebied van AI	31
4.4 Reflectie op de empirische verwachtingen	32
5. Conclusie & Discussie	34
5.1 Conclusie	34
5.2 Tekortkomingen onderzoek	35
5.3 Aanbevelingen vervolgonderzoek	35
Literatuurlijst	36

Bijlagen	39
1. Tabel: Schematische weergave kenmerken Acosta et al (2022)	39
2. Topiclijsten & interviewprotocol	40
2.1 Interviewprotocol	40
2.2 Topiclijst experts op het gebied van de governance van AI	40
2.3 Topiclijst experts op het gebied van het adviesstelsel en adviesorganen	41
3. Specificatie respondenten	42
4. Specificatie documenten	43

1. Inleiding

1.1 Aanleiding

Het is vrijdag 15 januari 2021 als minister-president Mark Rutte in een persconferentie aankondigt dat hij het ontslag van het voltallige kabinet heeft aangeboden aan de koning. De oorzaak van het vallen van kabinet Rutte III is het rapport *Ongekend onrecht*, dat een maand eerder werd gepubliceerd door de Parlementaire ondervragingscommissie Kinderopvangtoeslag (NOS, 2021a). In haar rapport is de ondervragingscommissie spijkerhard, zo stelt zij dat er grondbeginselen van de rechtsstaat zijn geschonden en dat alle drie de machten zich dat aan mogen rekenen (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020). De vragen: *hoe heeft het zo ver kunnen komen?* en *hoe kon het zo mis gaan?* domineren weken, maanden en zelfs jaren later nog steeds in de media (oa RTL, 2020. NOS, 2021b. NOS, 2022).

Jaren later is het toeslagenschandaal hét toonbeeld van wat er allemaal mis gaat bij de rijksoverheid. Zo kaart het institutionele racisme binnen de belastingdienst aan (RTL, 2022. Van Rij, 2022). Het toont aan dat de informatievoorziening vanuit de rijksoverheid niet op orde is (Parlementaire ondervragingscommissie Kinderopvangtoeslag, 2020). Tot slot laat het zien hoe het gebruik van een zelflerend algoritme bij de rijksoverheid verschrikkelijk fout kan gaan (Volkskrant, 2021). Zodoende laat het rapport van de ondervragingscommissie nog meer vragen rijzen: *is er nog meer racisme binnen de overheid? kan de Tweede Kamer haar werk nog doen met deze gebrekkige informatievoorziening? laat de overheid zich niet adviseren wanneer zij AI of zelflerende algoritmen gebruikt?* Hoewel al die vragen zich lenen voor tal van onderzoeken, zal deze thesis de vraag rondom advies op AI-gebruik door de overheid behandelen. Het onderwerp vergaart zo ook meer dan alleen maatschappelijke belangstelling, zo is het ook terug te vinden in de academische literatuur (oa. Giest & Klievink, 2022).

Men kan anno 2024 niet meer om het gebruik van kunstmatige intelligentie heen, zo ook bij de rijksoverheid. Het toeslagenschandaal laat zien hoe pijnlijk het gebruik van AI fout kan gaan, maar dat doemscenario gaat gelukkig niet op voor alle AI-toepassingen die gebruikt worden. Het gebruik van AI-toepassingen brengt veel kansen en mogelijkheden met zich mee, in de positieve zin van het woord. Zo kan het ingezet worden binnen de gezondheidszorg en maakt het het doen van inkomstenbelasting een stuk makkelijker (Henman, 2020, pp. 210-213). Tevens is er in de academische discussie over hoe een overheid om dient te gaan met AI veel ruimte voor zowel de voor- als nadelen van dit AI-gebruik. Daar worden dan vaak governance-modellen of frameworks aan gekoppeld (oa. Almeida & Gasser, 2017; Sturm, Weyer & Wirtz, 2020). De behoefte aan dit soort modellen wordt getekend door de hoeveelheid wet- en regelgeving die men tracht te creëren voor AI, bijvoorbeeld de Europese AI-act. Die behoefte naar institutionele sturing op AI blijft groeien, zo stuurde Von der Leyen vorig jaar een brief waarin zij een Europese AI-raad aankondigde. Ook in Nederland is deze behoefte groeiende, zo werd de motie van de leden

Dekker-Abdulaziz en Rajkowski die het kabinet oproept om een Nederlandse AI-raad te onderzoeken aangenomen. Daarmee groeit ook de behoefte naar (academisch) onderzoek op het snijvlak van advisering en AI. Feit is echter dat de literatuur over hoe overheden zich kunnen laten adviseren over het gebruik van AI nauwelijks bestaat. Deze thesis beoogt een begin te zijn om die mismatch te ondervangen, en daarmee het snijvlak van advisering en AI (hoger) op de academische agenda te zetten.

1.2 Probleemstelling (vraag- en doelstelling)

Er is academische literatuur over advisering aan de overheid in allerlei vormen, van factoren die meespelen bij het inrichten van een dergelijk orgaan tot manieren van kennismanagement (oa. Acosta et al, 2022; Osborne, 2004). Die literatuur gaat zelfs een stap verder, naar beleidsbeïnvloeding (bijv. Aarts et al, 2022). Echter ontbreekt een academische blauwdruk wanneer men een adviesorgaan wil inrichten; er zijn weinig tot geen peer-reviewed artikelen over wat nou wel en niet effectief is op dit vlak. Daarnaast is er literatuur waarin een (academisch) fundament wordt gelegd voor wet- en regelgeving van AI. Die literatuur varieert van de vraag of er überhaupt aparte wet- en regelgeving voor AI moet zijn (Chauhan, 2022) tot complete denkrichtingen en uitgewerkte governancemodellen voor deze wet- en regelgeving (oa. Almeida & Gasser, 2017; Choi & Park, 2023). Zoals gesteld beoogt dit onderzoek deze literatuur te koppelen, zodat er een begin wordt gemaakt om dit kennisprobleem op te lossen. Deze thesis is daarmee een verkenning naar de mogelijkheden van een adviesorgaan op het gebied van AI. De hoofdvraag van dit onderzoek is: *Wat zijn de mogelijkheden voor het inrichten van een adviserend orgaan op het gebied van AI?*

Om een antwoord op die vraag te formuleren, zal er eerst een literatuurstudie gedaan worden. Daarin zullen de hierboven beschreven onderwerpen aan bod komen, zodat er een overzicht is over de verschillende facetten uit de literatuur waar dit onderzoek mee te maken heeft. Uit dit overzicht zal een topiclijst worden gedestilleerd om de verzamelde data mee te analyseren. Die analyse zal gebruikt worden om onderstaande empirische deelvragen te beantwoorden, die op hun beurt bij zullen dragen aan het beantwoorden van de hoofdvraag van deze thesis.

Empirische deelvragen:

1. *Hoe wordt huidige advisering op het gebied van AI vormgegeven?*
2. *Welke belangrijke factoren voor advisering op het gebied van AI identificeren experts?*
3. *Welke mogelijke inrichtingsvormen voor een adviserend orgaan op het gebied van AI identificeren experts?*

1.3 Maatschappelijke relevantie

De maatschappelijke relevantie van dit onderzoek is hoofdzakelijk tweeledig. Enerzijds moet een adviesorgaan op het gebied van AI-gebruik door een overheid de rechten van burgers en het algemeen belang beschermen en dienen. Daardoor zullen incidenten als het toeslagenschandaal als het goed is niet, maar in ieder geval een stuk minder snel plaatsvinden. Anderzijds vloeit uit het dienen en beschermen van mensenrechten en het algemeen belang ook de mogelijkheid voort dat de overheid AI-toepassingen in kan zetten die het leven van burgers makkelijker zullen maken.

1.4 Wetenschappelijke relevantie

Allereerst zit de wetenschappelijke relevantie hem in het opvullen van een gat in de literatuur. Er is simpelweg nauwelijks tot geen academische literatuur op het snijvlak van adviesorganen en AI. Daarmee is het tevens een aanvulling op de bestaande literatuur over zowel adviesorganen als governance van AI. Dat zit hem in het feit dat er middels documentanalyses en expertinterviews, vanuit een theoretische lens, conclusies zullen worden getrokken. Het betreft dus een mix van verschillende soorten data, daar waar veel wetenschappelijke literatuur over deze onderwerpen theoretische exercities óf casestudies zijn.

1.5 Leeswijzer

Dit onderzoek kent een logische opbouw. Na dit inleidende hoofdstuk volgt het theoretisch kader met daarin een duiding van de theoretische concepten. Vervolgens zullen de gemaakte keuzes met betrekking tot de onderzoeksmethoden worden verantwoord in de methodologie. Waarna in het resultatenhoofdstuk de resultaten van het onderzoek gepresenteerd zullen worden. Er wordt vervolgens afgesloten met de conclusie en discussie.

2. Theoretisch kader

2.1 Inleiding

Dat AI meer dan een simpel digitaal systeem is, en veel invloed heeft op de publieke sector en bureaucraten staat als een huis (oa. Giest & Klievink, 2022; Arnold & Barth, 1999; Bovens, Van Eck & Zouridis, 2019). Hoewel deze artikelen uit verschillende tijden afkomstig zijn, hebben zij één ding gemeen: ze beschrijven hoe een vorm van AI door de overheid gebruikt wordt bij haar dienstverlening of de uitvoering van beleid. Giest en Klievink gaan daarbij in hun artikel onder andere in op het schandaal rondom de kinderopvangtoeslag in Nederland. In hun beschrijving van deze casus komt naar voren hoe het gebruik van AI door de overheid op een pijnlijke manier fout kan gaan (Giest & Klievink, 2022, pp. 9-12).

Hoewel het toelagenschandaal een voorbeeld is van hoe erg AI-gebruik door een overheid mis kan gaan, liggen er ook mogelijkheden om een positieve impact te maken (Henman, 2020, pp. 210-213). Het feit dat het gaat om een systeemtechnologie die zowel positieve als negatieve effecten kan hebben of al heeft, maakt dat de roep om gerichte governance of regulering hard is. Een goed voorbeeld hiervan is de AI-act van de Europese Unie. Zoals gesteld in de probleemstelling is daar veel academische literatuur over. Artikelen vinden hun focus vooral op het creëren van conceptuele modellen of frameworks voor de governance van AI (oa. Almeida & Gasser, 2017; Sturm, Weyer & Wirtz, 2020; Choi & Park, 2023; Sharma, 2023). Of op de regulering van AI (oa. Jacobs en Simon, 2022; Singh Chauhan, 2022; Andrews, 2018).

Naast de roep om gerichte governance en regulering begint ook de vraag naar een adviserend orgaan op het gebied van AI steeds luider te worden. Niet alleen in Nederland (motie Dekker-abdulaziz Rajkowski), maar ook op internationaal niveau (brief van Von der Leyen). Zoals ook in de probleemstelling gesteld bestaat er een variatie aan academische literatuur waar het gaat over adviesorganen en advisering aan een overheid. Factoren waar men rekening mee dient te houden, manieren van kennismanagement en daadwerkelijke beleidsbeïnvloeding, het komt allemaal terug in de literatuur (oa. Acosta et al, 2022; Osborne, 2004; Aarts et al, 2022).

Zoals gesteld beoogt deze thesis een koppeling te maken tussen deze twee onderwerpen in de academische literatuur. Er is immers geen literatuur in de hoek van governance en regulering van AI wat een adviesorgaan of de rol daarvan onderzoekt. Tevens ontbreekt het in de literatuur over adviesorganen of advisering aan een overheid aan onderzoeken die over een dergelijk orgaan op het gebied van AI gaan, of die zich focussen op een soortgelijke nieuwe technologie. In dit theoretisch kader zullen de bovengenoemde onderwerpen behandeld worden, nadat de definitie van AI kort behandeld zal worden. Met deze exercitie wordt getracht een theoretische lens te creëren die enerzijds als leidraad zal fungeren bij de interviews, maar anderzijds gebruikt zal worden om de data mee te analyseren. Daarnaast zal

op basis van de behandelde literatuur enige verwachting worden uitgesproken betreffende de empirische resultaten.

2.2 AI

Binnen de academische literatuur zijn er tal van definities van AI, verschillend door het (vak)gebied waarin AI gebruikt wordt of door de mate van complexiteit van de AI. Binnen de bestuurskunde heeft men het vaak over (deels) geautomatiseerde besluitvormingsprocessen, waarbij een AI-toepassing bijvoorbeeld ingevulde formulieren voor toeslagen of de belastingaangifte analyseert (Giest & Klievink, 2022, p. 1). Echter zou er gesteld kunnen worden dat die definitie met de komst van generatieve AI, zoals Chat-GPT, al is ingehaald door de ontwikkelingen. Het feit dat die ontwikkelingen zo snel gaan, maakt ook deel uit van de reden dat er geen echte academische consensus is over een definitie van AI. Wel identificeert de literatuur verschillende vormen van AI zoals machine learning, waarbij de AI-toepassing zelflerend vermogen heeft (Henman, 2020, p. 210). Sommige academici beschrijven de volgende stap, waarbij het voor de mens niet meer duidelijk is op basis waarvan AI een bepaalde beslissing neemt, als black-box toepassingen (Almeida & Gasser, 2017). Kortom, er zijn verschillende definities van AI en er zijn verschillende soorten AI.

Voor dit onderzoek zal AI in de breedst mogelijke zin van het woord gebruikt worden, waarbij wanneer nodig toegelicht zal worden of het gaat om bijvoorbeeld een analyserende toepassing, simpele rekenregels of machine learning oid. De reden hiervoor is tweeledig: allereerst omdat een algemeen aanvaarde definitie van AI in de academische wereld lastig te formuleren en te vinden is. Ten tweede omdat deze benadering alle mogelijke ruimte voor het onderzoek biedt, er wordt op deze manier dus niet aan de voorkant al afgebakend op wat voor soort AI-toepassingen adviezen gegeven zouden moeten worden. Dit brengt echter wel het risico met zich mee dat onduidelijk wordt waar nou eigenlijk over wordt geschreven en wat de relatie tussen AI en de overheid is. Om dat te ondervangen zal eerst worden stilgestaan bij de reden waarom de relatie tussen AI en de overheid een uitdagende is en waarom (gerichte) regulering nodig is.

2.3 Regulering en Governance van AI

Zoals eerder gesteld is er geen gebrek aan academische literatuur op het gebied van governance en regulering van AI. In de introductie is al aangestipt waar deze literatuur vaak over gaat: het creëren van een (integraal) governance model voor AI in de publieke sector en/of regulering van AI. Daarbij gaat de meest recente literatuur over regulering over de AI-act van de EU (oa. Jacobs & Simon, 2022), en worden verschillende theoretische blikken geworpen op het *hoe* en *waarom*. Chauhan probeert in zijn artikel vooral een antwoord te formuleren op de vragen waarom de relatie tussen AI en de overheid ingewikkeld is en waarom regulering van AI nodig zou zijn (Chauhan, 2022). In zijn artikel doet hij dit vanuit vier *imperatives*, dat zijn *uniqueness*, *protection*, *institutional* & *aspirational*. Om de *waarom-vraag* te voorzien van een theoretische basis zal onder andere dit artikel gebruikt

worden. Vervolgens zullen verschillende aspecten van zowel governance-modellen als perspectieven op regulering behandeld worden.

2.3.1 Waarom regulering?

In zijn artikel bespreekt Chauhan dus de noodzaak van regulering van AI langs vier hoofdlijnen, waarmee hij ingaat op voor- en tegenargumenten. Allereerst *uniqueness*, daar gaat Chauhan in op de vraag of AI wel uniek genoeg is voor eigen regulering. Tegenstanders betogen dat techniek in zijn geheel en daarmee dus ook AI geen speciale aandacht behoeven van het recht (Easterbrook, 1996 in Chauhan, 2022, p. 103). De implicatie in dat gedachtegoed is dat de huidige wet- en regelgeving afdoende zijn. Daar tegenover zet Chauhan het argument van Chesterman, zijnde dat er een drietal kenmerken zijn die maken dat AI wel degelijk eigen regulering behoeft. Dat zijn de snelheid waarmee AI beslissingen of acties kan (onder)nemen, de autonome aard van de beslissingen/acties en het feit dat het lastig is om beslissingen genomen door AI systemen/toepassingen te interpreteren, zeker wanneer sprake is van machine learning (Chesterman, 2021 in Chauhan, 2022, pp. 103-104). Tot slot betoogt Chauhan dat de situatie zich voor kan doen dat er tegen een beslissing die genomen is door AI geprocedeerd kan worden bij de rechter, en dus een duidelijk verantwoordelijk persoon ontbreekt. Dit toont volgens hem de noodzaak van aparte wet- en regelgeving ten behoeve van de regulering van AI aan (Chauhan, 2022, p. 104).

Naast de unieke eigenschappen gaat Chauhan ook in op *protection, institutional en aspirational imperatives* die volgens hem laten zien waarom AI gereguleerd moet worden. De belangrijkste constatering uit die drie zijn dat (nieuwe) technologieën altijd op een bepaald punt in hun ontwikkeling gereguleerd dienen te worden (en dat dit ook gebeurt). Verder gaat de auteur in op het feit dat er bij dit soort technologieën, die een grote impact hebben op de samenleving, sprake zou moeten zijn van een bepaalde legitimiteit van de geldende regels. Met andere woorden, dat de regulering opgepakt zou moeten worden door een centrale statelijke actor die een democratisch legitieme basis heeft om dit soort regulering voor de hele staat in te voeren. Dat is bijvoorbeeld niet het geval bij een bedrijf of NGO dat een normenkader voor intern gebruik van AI op heeft gesteld. Tot slot gaat Chauhan in op het feit dat veel van de huidige wetenschappelijke literatuur over AI ingaat op de gevaren, op de dingen die fout kunnen gaan. Dat terwijl er volgens hem ook heel veel kansen liggen bij het gebruik van AI door de overheid, mits er dus adequate regulering is (Chauhan, 2022, pp. 104-110).

In zijn conclusie neemt Chauhan een stelling in, waarmee hij ook andere academici aan zijn zijde vindt. Hij stelt namelijk dat een belangrijke factor in het succes van de regulering de preventieve en proactieve aard daarvan is (Chauhan, 2022, pp. 120-122). Onder andere Nordström onderstreept het belang van deze aard, zij stelt dat voor een succesvolle regulering in de publieke sector preventieve en een proactieve regulering essentieel zijn (Nordström, 2022). Echter zijn er ook academici die juist hameren op het belang van controle (en eventueel sancties) achteraf, waarmee er dus een meer reactieve vorm van regulering wordt voorgesteld. Interessant daarbij is dat Baum et al daarvoor redenen geven die in lijn zijn met

de argumentatie van Chauhan en Nordström (Baum et al, 2023). Er is dus sprake van een consensus over het feit dat er een vorm van regulering moet komen, maar hoe die vorm er uit ziet en hoe de manier van reguleren zou moeten zijn, is men het (nog) niet eens.

2.3.2 Governancemodellen

In de wetenschappelijke literatuur wordt veel geschreven over governancemodellen voor de governance van AI. Daarbinnen valt een onderscheid te maken in grofweg twee categorieën: modellen voor private organisaties en modellen voor de publieke sector. Voor dit onderzoek zal alleen gekeken worden naar de modellen die gecreëerd zijn voor de publieke sector, omdat advisering over AI aan een overheid de focus van dit onderzoek is. Binnen die literatuur is er veel onderscheid in de soorten modellen, de een betoogt een gelaagd model (Almeida & Gasser, 2017; Sturm, Weyerer & Wirtz, 2020) daar waar anderen een integraal framework optuigen (oa. Sigfrids et al, 2022; Choi & Park, 2023; Sharma, 2023). In deze paragraaf zal gekeken worden naar de theoretische argumentatie achter de modellen. Het antwoord op de onderliggende *waarom*-vraag is al behandeld in deze thesis, maar wordt goed aangevuld door Almeida, Mendes en Doneda (2023). Zij stellen namelijk dat een governancemodel of framework nodig is om enerzijds de publieke veiligheid en mensenrechten te beschermen, en anderzijds om een flexibele en innovatieve omgeving te creëren zodat we niet blind zijn voor de kansen van AI. Het antwoord op de vraag waarom bepaalde elementen al dan niet zouden moeten worden meegenomen wanneer men nadenkt over de governance van AI zal hier aan bod komen.

In het gelaagde model van Almeida en Gasser (2017), staat centraal dat er een kosten-batenanalyse gemaakt moet kunnen worden bij het gebruik van AI. Hierin faciliteert hun model dan ook. Daarbij onderstrepen de auteurs tevens het belang van het vinden van normatieve consensus onder de betrokken stakeholders, met andere woorden dat de stakeholders het eens zijn over een normatieve basis waarop het governancemodel gebouwd dient te worden. De lagen zijn (van onder naar boven): het technische fundament, de ethische laag en de maatschappelijke- en gerechtslaag. De logica in dit model zit hem volgens de auteurs in het feit dat de ethische en maatschappelijke- en gerechtslaag met elkaar interacteren, op basis van het technische fundament. Er wordt dus gedacht vanuit dat technische fundament naar bijvoorbeeld criteria of ethische principes die op hun beurt uit zouden moeten monden in normen en wet- en regelgeving (Almeida & Gasser, 2017, pp. 59-61). Dit model voor de totstandkoming van wet- en regelgeving biedt ruimte voor een mogelijk adviesorgaan, bijvoorbeeld op de ethische principes. Daarnaast zou deze visie of manier van werken ook kunnen worden overgenomen door een adviserend orgaan, door bijvoorbeeld in verschillende samenstellingen per laag te adviseren.

In tegenstelling tot voorgaande auteurs zetten Sturm, Weyerer en Wirtz (2020) de uitdagingen die AI-gebruik met zich meebrengt centraal in de argumentatie voor hun governance framework, en gebruiken zij de *regulation theory*. Daarbij identificeren zij wel dezelfde zaken als Almeida en Gasser (2017). Zo verdelen zij de uitdagingen in drie categorieën: maatschappelijke, ethische en uitdagingen omtrent wetgeving en regulering (Sturm, Weyerer

& Wirtz, 2020, pp. 820-822). Zij noemen dit de uitdagingen-laag in hun framework en hoewel deze uitdagingen in hun argumentatie centraal staan, staan zij dat niet in het framework zelf. Zij zetten namelijk de *AI Regulation Process Layer* centraal in dit framework. Binnen die laag spelen framing, kosten- batenanalyse, evaluatie en risicomanagement ieder een eigen rol in het proces. Binnen deze laag zou ruimte kunnen zijn voor een adviserend orgaan, bijvoorbeeld betreffende risicomanagement of de kosten-batenanalyse. Daar waar Almeida en Gasser het faciliteren van deze analyse centraal zetten in hun model, wordt het hier dus geïntegreerd in het reguleringsproces. De argumentatie daarvoor vinden zij echter in het artikel van Almeida en Gasser, zo stellen de auteurs dat het proces op deze manier zorgvuldig, voor alle uitdagingen in de toekomst, het hoofd kan bieden (Almeida & Gasser, in Sturm, Weyerer & Wirtz, 2020, p. 825). De overige lagen zijn de *AI Applications and Technology Layer*, de *Public AI Policy Layer* en de *Collaborative AI Governance Layer*. Echter, wat de centrale laag het meest relevant maakt is het feit dat de auteurs als een van de eersten ingaan op hoe het proces rondom de totstandkoming van wet- en regelgeving op het gebied van AI eruit zou kunnen zien. De manier waarop dit proces wordt ingericht biedt zoals gesteld ruimte voor een adviserend orgaan. De vraag die hierdoor alleen wel wordt opgeworpen is of de relatie hier zo is dat het orgaan binnen een bepaalde laag functioneert, of dat zij overkoepelend te werk dient te gaan. Deze vraag vloeit voort uit het verschil in wat er precies centraal staat in de modellen.

Ook zijn er auteurs die de stelling wat scherper neerzetten. Zo geven Choi en Park (2023) hun artikel een pakkende titel: *“To govern or to be governed: an integrated framework for AI governance in the public sector”*. Vervolgens trekken zij het doemscenario in hun introductie verder door, door vergelijkingen te trekken met de wereld zoals beschreven door George Orwell in het boek 1984 (Choi & Park, 2023, p. 1059). Desalniettemin brengen ze een interessante blik op de governance van AI. Zij borduren met hun framework door op onder andere de conceptuele ideeën van Almeida en Gasser en Sturm, Weyerer en Wirtz (Choi & Park, 2023, p. 1062). Het perspectief wat zij toevoegen aan de academische discussie is dat van IT-governance. Zij focussen daarmee veel meer op IT-activiteiten die ten grondslag liggen van IT-innovaties als AI, zoals IT-infrastructuur management, IT-gebruik management en product management. Zij koppelen deze theorie aan AI, samen met de technische (ontwikkelings)fases van een AI-toepassing, het type beslissing (bijvoorbeeld over ethische principes voor AI) en welk soort governance (centraal, decentraal, etc.) daar bij zou horen. De auteurs passen dit toe op de casus Zuid-Korea, wat daaruit volgt is een theoretische en empirische basis voor de toepassing van deze matrix (Choi & Park, 2023, pp. 1063-1069). De focus op juist de technische kant van AI en de IT-zaken daaromheen, maakt dat dit weer een ander perspectief biedt op de governance van AI en de relatie tussen AI en de overheid. Hierbinnen is tevens plek voor een adviserend orgaan, maar nog interessanter is de visie die betoogd wordt. Namelijk een andere relatie tussen AI en de overheid en het feit dat er verschillende lagen in die relatie zitten. Dit verschil in inzicht biedt de mogelijkheid om verder te kijken dan bijvoorbeeld een centrale overheid en de maatschappelijke impact van een AI-toepassing, wanneer er een adviserend orgaan zou worden ingesteld.

Ondanks de nuances en perspectieven die verschillende auteurs aanbrengen in hun bijdragen aan de academische discussie op het gebied van regulering en governance van AI, is er een heldere rode draad door alle artikelen te vinden. In de behandelde literatuur, maar ook door andere auteurs (oa. Sharma, 2023; Birkstedt et al, 2023; Kuziemski & Misuraca, 2020) wordt de noodzaak van verder onderzoek naar de governance van AI onderstreept. Daarbij sporen sommige aan om voort te bouwen op huidige frameworks of modellen, om zo tot een consensus te komen die heel breed ingezet kan worden (Kuziemski & Misuraca, 2020, pp. 9-12). Terwijl anderen aangeven dat pionieren en theoretische gedachte spinsels de wetenschappelijke discussie verder zullen brengen (oa. Choi & Park, 2023; Almeida, Mendes & Doneda, 2023). Deze thesis zal een combinatie van beide zijn, omdat zij enerzijds voortbouwt op hetgeen reeds beschreven is in de literatuur. Anderzijds zal dat voortbouwen een verkenning zijn naar iets waar nog vrij weinig (academische) kennis over is: een adviesorgaan op het gebied van AI.

2.4 Adviesorganen

Na het behandelen van de inhoud over AI, is het noodzakelijk om ook de andere kant van dit onderzoek te voorzien van een theoretische basis. In deze paragraaf wordt gekeken naar de theoretische beschrijvingen van verschillende soorten adviesorganen, de onderliggende manieren van kennismanagement en de uiteindelijke beleidsbeïnvloeding. Doel hiervan is de rationale achter verschillende soorten organen te achterhalen en begrijpen, zodat elementen daarvan gebruikt kunnen worden bij het analyseren van de onderzoeksdata en het formuleren van de conclusies.

2.4.1 Theoretische beschrijvingen

Om te kunnen begrijpen wat er speelt in de wetenschappelijke discussie over adviesorganen, is het goed om te starten bij de mogelijkheden. Dit is nodig zodat de kennis over wat een adviesorgaan zou moeten kunnen, weten en doen in kaart gebracht kan worden. Het uiteindelijke doel daarvan is dat deze kennis meegenomen kan worden in de dataverzameling en analyse, ten behoeve van het beantwoorden van de hoofdvraag. Met andere woorden, welke manieren zijn er om een adviesorgaan in te richten? Om dat in kaart te brengen hebben Acosta et al (2022) twaalf verschillende kenmerken van adviesorganen op papier gezet en voorzien van allerlei theoretische facetten. Alvorens dit behandeld zal worden is het goed om een notie van de auteurs over te brengen. Zij stellen namelijk dat het bijna onmogelijk is om de impact en effectiviteit van verschillende soorten adviesorganen te meten, en dat dit daarom ook niet gebeurd is in peer-reviewed artikelen (Acosta et al, 2022, pp. 17-18). Van de twaalf kenmerken zijn er zeven relevant voor dit onderzoek, deze worden schematisch weergegeven in de tabel in bijlage 1. De auteurs concluderen hun overzicht met de stelling dat het feit dat onderzoek naar de effectiviteit van bepaalde adviesorganen er niet is, maakt dat hun overzicht niet als handboek of blauwdruk gebruikt kan worden (Acosta et al, 2022, pp. 20-21). Dit is dus een overzicht van verschillende keuzes die gemaakt kunnen worden, aan de hand van bepaalde kenmerken van het adviesorgaan en de context waarin zij acteert. Deze kenmerken

en mogelijke keuzes worden meegenomen in de dataverzameling, ter houvast en om een gesprek over deze mogelijkheden te starten.

Een diepere laag in de academische literatuur over adviesorganen wordt geboden door oa. Crowley en Head (2017). Zij bespreken de manier waarop adviesraden gevuld met experts acteren in de huidige politiek-bestuurlijke context. Crowley en Head suggereren dat adviesorganen bruikbaar en misschien noodzakelijk zijn op thema's waar een mening niet voldoende is om het algemeen belang te beschermen en te dienen (Crowley & Head, 2017, p. 16). Voor dit onderzoek is dat een relevante constatering, omdat het aannemelijk is dat dit ook voor AI geldt. Het lastige daarbij is dat deze adviesorganen niet alleen meer gelegitimeerd dienen te worden door de politiek-bestuurlijke spelers, maar door een steeds breder publiek, als zij effectief willen handelen. Dit noemen de auteurs een onderdeel van de externalisatie van adviesorganen (Crowley & Head, 2017, pp. 4-7). Ter illustratie: dit werd goed zichtbaar tijdens de Covid-19 pandemie, waar het Nederlandse OMT en haar leden onderwerp werden van de discussie omdat haar adviezen niet altijd even breed gedragen werden door de Nederlandse samenleving (Aarts et al, 2022). Tevens trachten de auteurs verschillende inzichten over het soort adviesorgaan en verschillende soorten adviezen te integreren, waarmee zij een theoretische basis leggen voor wat voor soort orgaan wellicht beter bij een bepaalde soort adviestaak past. Daarbij houden zij ook rekening met de factoren die de externalisatie met zich meebrengt, waardoor er eigenlijk een soort toolbox ontstaat voor adviesorganen (Crowley & Head, 2017). Hetgeen Acosta et al (2022) nuanceren door te stellen dat er geen onderzoek gedaan is naar de effectiviteit van verschillende soorten adviesorganen. Echter, de toolbox van Crowley en Head (2017) kan een rol spelen in dit onderzoek, omdat ze niet sec kijken naar de interne structuren van adviesorganen maar ook rekening houden met zaken als de publieke opinie over een adviesorgaan, de thema's waarover geadviseerd wordt en het algemeen belang.

2.4.2 Kennismanagement

Tot slot kijkt Osborne (2004) naar kennismanagement bij adviesorganen, een vitaal element wat meegenomen dient te worden bij bijvoorbeeld de selectie en benoemingsprocedure van leden van een adviesorgaan en de op te zetten processen binnen een adviesorgaan (Osborne, 2004). Een notie die hierbij gegeven moet worden is dat Osborne zijn artikel schreef in de context van een adviesraad in Australië, gespecialiseerd in familierecht. Dat is iets wezenlijk anders dan AI, desalniettemin levert de input over kennismanagement bij adviesorganen relevante inzichten op voor dit onderzoek. Een van die zaken is het opstellen van een kennismanagement strategie, waarbij een tweetal vragen centraal staan: *hoe wordt kennis in de raad bewaard en overgebracht?* en *hoe interacteren verschillende soorten kennis in de raad?* (Osborne, 2004, p. 46). Om aan te tonen waar de relevantie voor dit onderzoek zit, kan er gedacht worden aan hoe kennis over ethische bezwaren over een AI-toepassing interacteert met de technische kennis over diezelfde toepassing. Wat prevaleert en waarom? Vervolgens gaat Osborne in op de rol van het adviesorgaan zelf bij het creëren van kennis. Met andere woorden, gaat er meer gebruikgemaakt worden van bestaande kennis of is het de bedoeling dat het orgaan zelf kennis gaat creëren over bepaalde onderwerpen (Osborne, 2004, p. 47).

Voor dit onderzoek is dat een relevante stap bij het maken van zo'n strategie, omdat er sprake is van een techniek waarvan men zeker weet dat de ontwikkelingen voorlopig niet zullen stoppen. Osborne concludeert met de stelling dat kennismanagement een essentieel onderdeel van het proces van een adviesorgaan is, en dat het nooit af is. Het is een incrementeel proces waar (periodieke) bijsturing een inherent onderdeel van is (Osborne, 2004, pp. 50-51). Hetgeen relevant is voor dit onderzoek, omdat enerzijds de ontwikkelingen rondom AI zullen vragen om bijsturing en anderzijds de ervaringen van het werken met dit nieuwe adviesorgaan dat wellicht zullen doen.

In deze paragraaf is getracht een theoretische basis te leggen voor de mogelijkheden betreffende het inrichten van een adviesorgaan. Daarvoor is een deel van de academische discussie over adviesorganen en hun kennismanagement gebruikt. Het enige wat expliciet ontbreekt in deze academische discussie is een empirische stap waarmee de effectiviteit van bepaalde vormen of keuzes aangetoond kan worden.

2.5 Theoretische lens & empirische verwachtingen

De theoretische lens die uit dit hoofdstuk wordt gedestilleerd wordt weergegeven in 3.4 *Operationalisatie van theoretische concepten*. Daarnaast zijn de topiclijsten die gebruikt zullen worden bij het afnemen van de interviews opgenomen in bijlage 2. Er is bij deze thesis dus geen sprake van een geconceptualiseerd model wat gebruikt gaat worden bij de dataverzameling en analyse.

Vervolgens de empirische verwachtingen. Vanuit de behandelde literatuur kunnen een aantal verwachtingen worden uitgesproken ten aanzien van de te verzamelen empirische data. Allereerst de verwachtingen met betrekking tot de governance van AI. Te beginnen met de vraag of er wel aparte wet- en regelgeving en advisering nodig is op het gebied van AI. De verwachting is dat de data zal uitwijzen dat aparte wet- en regelgeving en advisering op het gebied van AI wenselijk dan wel noodzakelijk is. Ten tweede kan vanuit de theorie de verwachting worden uitgesproken dat proces betreffende de totstandkoming van wet- en regelgeving op het gebied van AI centraal staat. Waarbij een adviserend orgaan, vanuit verschillende visies (bijvoorbeeld technisch, maatschappelijk, etc.) een rol kan spelen. Daarnaast is de verwachting dat een dergelijk orgaan overkoepelend te werk dient te gaan, dit vanwege het feit dat AI op zoveel verschillende domeinen en overheidslagen impact heeft. Tot slot de factoren die als belangrijk worden gezien op het gebied van advisering over AI. De verwachting is dat daar in ieder geval een diversiteit aan kennis zal worden aangedragen. Dat omdat AI een systeemtechnologie is die ontzettend veel vlakken raakt, zoals ook blijkt uit de literatuur. Denk hierbij aan technische experts, maar bijvoorbeeld ook aan filosofen en juristen.

Vervolgens de verwachtingen die slaan op de mogelijke inrichtingskeuzes van een dergelijk adviesorgaan. Hierbij dient eerst de notie gegeven te worden dat er geen artikelen gevonden zijn die peer-reviewed zijn en duidelijk laten zien welke manier van inrichting effectief is en waarom. Er zijn in dit geval hoofdzakelijk een drietal mogelijkheden gedestilleerd uit de theorie, voor de belangrijkste keuze: er komt geen adviesorgaan op het gebied van AI, er komt een adviesorgaan op het gebied van AI of het onderwerp AI wordt ondergebracht bij een bestaande adviesraad. De verwachting is dat de keuze op een van de laatste twee zal vallen, waarbij de doorslaggevende factor voor de keuze tussen die twee de beoogde rol van het adviesorgaan zal zijn. Daarnaast is de benoeming van leden en de integriteit van een orgaan, en daarmee de autoriteit van het orgaan en haar leden van levensbelang voor de acceptatie van haar adviezen. Afsluitend de organisatie van het kennismanagement, vanuit de literatuur kan de verwachting worden uitgesproken dat de data het belang van incrementaliteit laat zien. Dit omdat de techniek constant in beweging is en periodieke bijsturing dus gewenst is.

3. Methoden & Technieken

3.1 Inleiding

Dit hoofdstuk licht de methoden en technieken toe die gebruikt worden om de deelvragen, en daarmee de hoofdvraag van deze master thesis te beantwoorden. Allereerst zal de onderzoeksstrategie besproken worden. Vervolgens zullen de dataverzameling en de data-analyse beschreven en onderbouwd worden. Daarna zal de vertaalslag van theorie naar empirie gemaakt worden in de operationalisatie. Tot slot zal er een kritische blik gegeven worden op de validiteit en betrouwbaarheid van dit onderzoek.

3.2 Onderzoeksstrategie

Het doel van dit onderzoek is enerzijds beschrijvend, anderzijds verkennend. Beschrijvend omdat de bestaande literatuur over adviesorganen, advisering aan een overheid en governance van AI geraadpleegd wordt om de relevante inzichten over dit thema (en daarmee dit onderzoek) te beschrijven. Het verkennende doel van dit onderzoek zit hem in het feit dat er aan de hand van de data-analyse mogelijkheden en belangrijke zaken in kaart worden gebracht, met betrekking tot een adviserend orgaan op het gebied van AI. In deze thesis is gekozen voor een kwalitatieve benadering, zodat er een diepgaand begrip over het onderzoeksonderwerp verkregen kan worden.

Er is dus gekozen voor een kwalitatieve benadering, omdat de onderzoeker daarmee de mogelijkheid krijgt om het onderwerp diepgaander te analyseren (Bryman, 2016). Daarnaast is kwalitatief onderzoek een goed instrument om onderzoek te doen naar zaken waarnaar weinig onderzoek is gedaan (Boeije & Blijenbergh, 2019, p. 30). Die combinatie maakt dat deze onderzoeksmethode uiterst geschikt is voor dit onderzoek. Het doen van een diepgaande analyse maakt dat er tijdens de dataverzameling en data-analyse betekenisvolle inzichten kunnen worden opgedaan, die verder gaan dan een bepaalde statistische waarde. Daarbij komt dat dit onderzoek deels verkennend is, er is dus weinig onderzoek gedaan naar dit onderwerp. Daarvoor is een kwalitatieve benadering zoals gezegd geschikt, bij deze thesis zal dat expliciet terugkomen bij de interviews. De kwalitatieve benadering biedt namelijk ruimte voor nieuwe ideeën en inzichten die de respondenten kunnen opperen. De data zal naast semi-gestructureerde interviews ook verzameld worden middels een documentanalyse. Hierover leest u meer in paragraaf 3.3 *Dataverzameling en data-analyse*.

3.3 Dataverzameling en data-analyse

De methoden van dataverzameling en -analyse die in dit onderzoek zijn gebruikt, worden in deze paragraaf verantwoord. Zoals al eerder gesteld zal er in dit onderzoek gebruik worden gemaakt van meerdere manieren om data te verzamelen. Hierdoor ontstaat een completer beeld van de te onderzoeken materie, dit heeft op zijn beurt een positief effect op de validiteit van het onderzoek (Boeije & Blijenbergh, 2019, p. 160) In deze paragraaf worden de te gebruiken methoden dus verantwoord. Tot slot wordt er in de laatste subparagraaf de data-analyse beschreven.

3.3.1 Semi-gestructureerde interviews

Het gebruik van semi-gestructureerde interviews zorgt ervoor dat de interviewer het verloop van het interview op een goede manier kan voorbereiden, maar dat er ook ruimte blijft om te kunnen doorvragen als er een interessant thema of iets onduidelijks wordt genoemd door de respondent (Boeije & Blijenbergh, 2019, pp. 74-75). Deze manier van interviewen voorziet dus in het beste van beide werelden. De interviews zullen worden ingezet ten behoeve van het beantwoorden van de hoofdvraag, doordat zij de empirische deelvragen beantwoorden. Dit zal gedaan worden met hun inzichten op het gebied van de governance van AI en adviesorganen. Daarmee dragen zij bij aan het verkennende deel van het onderzoek, de data die deze interviews genereren zullen vanuit de theoretische lens bekeken worden om zo tot een uiteindelijke inventarisatie van de mogelijkheden en belangrijke zaken te komen.

De respondenten zijn hoofdzakelijk geselecteerd aan de hand van hun expertise en hun beschikbaarheid. Voor dit onderzoek zijn een drietal expertisegebieden geselecteerd: governance van AI, advisering aan een overheid en het Nederlandse adviesstelsel. Met deze drie expertisegebieden is getracht om op hoofdlijnen de onderwerpen die met de dataverzameling verzameld worden af te bakenen, maar tegelijk ruimte voor variatie te behouden. In totaal zijn hiervoor negentien unieke experts benaderd voor deelname aan dit onderzoek. Daarvan hebben er zeven daadwerkelijk zelf geantwoord. Van deze zeven gaf een drietal aan deel te willen nemen aan dit onderzoek. In de tabel in bijlage 3 vindt u een verdere specificatie van de geselecteerde respondenten. Om de interviews af te nemen is allereerst vanuit de bestudeerde literatuur een topiclijst opgesteld, deze vindt u terug in bijlage 2. In deze bijlage staan verschillende versies van de topiclijst, afgestemd op de respondent, zodat de expertise van de respondent tot haar recht komt. Ook zal het interviewprotocol in deze bijlage te vinden zijn, zodat te zien is hoe de respondenten op de hoogte zijn gebracht van hun rechten en hoe hun toestemming is verkregen.

3.3.2 Documentanalyse

Zoals gesteld zal het uitvoeren van een documentanalyse ook onderdeel zijn van dit onderzoek. Documenten kunnen op verschillende wijzen een bijdrage leveren aan een onderzoek (Bryman, 2016, pp. 546 - 567), in dit geval voorziet het in feiten over het onderzoeksonderwerp en zorgt het daarmee voor triangulatie. Daarnaast zullen verschillende documenten ook gebruikt worden om inzichten en meningen over het onderzoeksonderwerp te vergaren. Denk hierbij aan stukken die over AI gaan, maar ook aan relevante wetgeving of documenten die door respondenten zijn aangedragen. Het doel van de documentanalyse is tweeledig. Enerzijds geeft het meer context bij het onderwerp, door relevante en aanverwante documenten te analyseren. Anderzijds draagt het bij aan het beantwoorden van de empirische deelvragen, door in te gaan op wat er nu al bekend is over advisering op het gebied van AI. Zo draagt het bij aan zowel het beschrijvende als verkennende aspect van deze thesis.

Bryman heeft voor de bruikbaarheid van documenten een viertal criteria opgesteld: authenticiteit, geloofwaardigheid, representativiteit en betekenis (Bryman, 2016, p. 546). De geselecteerde documenten zullen langs deze criteria worden gehaald. Daarnaast zijn de volgende criteria gebruikt om de documenten te selecteren: het hoofdonderwerp is AI, heeft met AI te maken of gaat over advisering aan een overheid. Door deze set aan criteria te gebruiken, is gewaarborgd dat er kwalitatief goede documenten geanalyseerd worden die ook inhoudelijk bijdragen aan het onderzoek. Een verdere uitwerking van de criteria per geselecteerd document vindt u in de tabel in bijlage 4.

3.3.3 Data-analyse

Om een antwoord op de hoofdvraag te kunnen geven is het van belang dat de data op een goede manier geanalyseerd zal worden. Allereerst zullen de interviews getranscribeerd worden¹, om deze data vervolgens aan thema's toe te kennen (Boeije, 2014, p. 93). Die thema's zijn terug te vinden onder *3.4 Operationalisatie theoretische concepten* en komen overeen met hetgeen in de topiclijsten is gezet. De geoperationaliseerde concepten zullen daardoor dus ook de leidraad zijn bij het analyseren van zowel de transcripties als de documenten. Door middel van deductief coderen zal vervolgens een antwoord op de empirische deelvragen, de hoofdvraag, en daarmee uiteindelijk een conclusie geformuleerd worden. Hierdoor zal dus binnen de totale puzzel (het theoretische plaatje), middels de empirische bevindingen, naar boven komen welke stukjes van deze puzzel meer of minder belangrijk gevonden worden (Boeije, 2014). Dit is dan ook het fundament van het verkennende deel van deze thesis.

¹ Hiervoor zijn twee AI-toepassingen gebruikt: de ingebouwde transcribeer-toepassing in MS-teams en Amberscript.

3.4 Operationalisatie theoretische concepten

Om de concepten uit het theoretisch kader meetbaar te maken, zullen deze geoperationaliseerd moeten worden. In tabel 1 is te zien op welke wijze dat is gedaan. De indicatoren zijn een praktisch bruikbare vertaling van de concepten, welke in de topiclijsten voor de interviews aan bod komen (bijlage 2).

Tabel 1

Operationalisatie

Theoretisch concept	Aspecten definitie	Indicatoren
Governance van AI	Manier waarop een overheid om zou moeten gaan met wet- en regelgeving op het gebied van AI.	<ul style="list-style-type: none"> - Relatie tussen AI en overheid - Noodzaak van aparte wet- en regelgeving - Soort wet- en regelgeving - Focus van wet- en regelgeving - Noodzaak van een adviesorgaan - Ruimte voor een adviesorgaan - Focus van een adviesorgaan op het gebied van AI - Functie van een adviesorgaan op het gebied van AI
Werking van adviesorganen in een politieke context	Factoren die de werking en effectiviteit van een adviserend orgaan kunnen beïnvloeden.	<ul style="list-style-type: none"> - Oordeel over wanneer een adviesorgaan goed functioneert - Functie van een adviesorgaan - Organisatie van kennismanagement - Organisatie van expertise - Mogelijke inrichtingskeuzes

3.5 Validiteit en betrouwbaarheid

Om de objectiviteit van een onderzoek te kunnen meten en te kunnen waarborgen zijn er in de wetenschap criteria opgesteld. In dit geval zijn de belangrijkste validiteit en betrouwbaarheid (Boeije & Blijenbergh, 2019; Bryman, 2016).

3.5.1 Validiteit

Om te beginnen met de interne validiteit, dat houdt in dat de getrokken conclusies daadwerkelijk voortvloeien uit het gedane onderzoek. Met andere woorden, wordt er gemeten wat er gemeten moet worden (Bryman, 2016, p. 385). Om dit zoveel mogelijk te garanderen zijn de concepten uit de literatuur geoperationaliseerd, zodat er een zeker mate van houvast is tijdens de interviews. Daarnaast zijn de respondenten op basis van de geformuleerde criteria geselecteerd. Hierdoor is geborgd dat er een match is tussen vraag en antwoord, wat de respondenten betreft. Tot slot zijn ook de gebruikte documenten voor de documentanalyses onderworpen aan een set criteria. Daarmee is getracht de interne validiteit op het niveau van

de academische standaard te krijgen. Vervolgens de externe validiteit, dit betreft de mate waarin de bevindingen gegeneraliseerd kunnen worden (Bryman, 2016, p. 385). Om dit zo goed als mogelijk te garanderen, is er bij het selecteren van de documenten en respondenten gebruikgemaakt van verschillende criteria. Hierdoor zouden andere documenten en respondenten die wel aan de criteria voldoen, als het goed is, tot dezelfde uitkomsten moeten leiden. Mits er gebruik is gemaakt van dezelfde analytische lens en dus hetzelfde theoretisch kader.

3.5.2 Betrouwbaarheid

Bij interne betrouwbaarheid heeft men het over het voorkomen van een tunnelvisie door bijvoorbeeld de mogelijkheid om te overleggen of sparren over de data en de analyses (Bryman, 2016, p. 384). Dit kan ultiem gewaarborgd worden door een onderzoek met meerdere onderzoekers uit te voeren, echter is dat bij dit onderzoek niet het geval. Wel was er de mogelijkheid om met de begeleidend hoogleraar en medestudenten uit de werkgroep te sparren over de data en de analyses.

Tot slot de externe betrouwbaarheid, die slaat op de mate van herhaalbaarheid van het onderzoek. Daarbij dient gelijk de kanttekening gemaakt te worden dat dit bij kwalitatieve onderzoeken als deze een heikel punt is (Bryman, 2016, p. 383). Dat komt doordat onderzoeken als deze hevig beïnvloed kunnen worden door de context waarin zij uitgevoerd worden. Desalniettemin is getracht de externe betrouwbaarheid te waarborgen door topiclijsten te gebruiken bij de interviews en door de respondenten en documenten te selecteren aan de hand van een set criteria. Hierdoor zouden de resultaten van dit onderzoek overeen moeten komen met een herhaling van dit onderzoek, als dezelfde selectiecriteria zijn gebruikt en de context niet hevig veranderd is.

4. Resultaten

In dit hoofdstuk zullen de empirische bevindingen gepresenteerd worden, gedestilleerd uit de vergaarde data. Zoals eerder gesteld is deze afkomstig uit interviews en documenten. Dit hoofdstuk is thematisch ingericht, dat wil zeggen dat de empirische bevindingen per thema zullen worden behandeld. Er is daarmee geen sprake van een behandeling per soort bron. Allereerst zal de governance van AI behandeld worden, waarna het functioneren en de effectiviteit van adviesorganen aan bod komen. Hierna zal de combinatie tussen deze twee gemaakt worden door advisering op het gebied van AI te behandelen. Dit hoofdstuk zal worden afgesloten met een reflectie op de eerder gestelde empirische verwachtingen.

4.1 Governance van AI

In deze paragraaf zullen de resultaten worden besproken die over de governance van AI gaan. De eerste indicator die daarbij een rol speelt is de kijk op de relatie tussen de overheid en AI. In dat opzicht waren de respondenten het over een ding in ieder geval eens met elkaar, AI is een technologie met een impact die dermate groot kan zijn dat de overheid er wel iets mee moet. Echter brengen zij wel verschillende nuances aan. Zo stelt R3 het volgende:

“Dat er een relatie tussen zou zijn impliceert dat er twee actoren zijn. Ik denk niet dat AI op dit moment al een eigen standing heeft, in de zin van dat het een rechtspersoonlijkheid moet krijgen. Dus ik bedoel dat er wel over nagedacht moet worden, maar daar zijn we nu echt nog niet.”

Daar staat de uitleg van R2 tegenover:

“Het beeld dat bij de overheid leeft over technologie, is vaak nog steeds een instrumenteel beeld. Dus technologie is een instrument, dat kunnen we goed gebruiken maar ook op een slechte manier. Maar het hele probleem is, zeker als het gaat om AI, dat technologie zelf ook iets doet met de mens. Dat is echt iets waar we denk ik vanaf moeten, willen we AI goed inbedden in de samenleving.”

Ook D2 biedt een uitleg over de relatie tussen de overheid en AI, althans zij schrijft voor hoe de overheid AI zou moeten zien. Daarin stelt de WRR dat de overheid (en de samenleving) AI moet zien als een systeemtechnologie. Daarmee doelen zij op een technologie die niet alleen in een specifiek vakgebied impact heeft, maar over de hele breedte. Zij vergelijkt dit met de komst van de auto en elektriciteit, die de wetgever dwingt de bredere effecten ervan op de samenleving te overwegen (WRR, 2021, p.17).

Deze vergelijking slaat gelijk een brug naar het volgende aspect, of er aparte wet- en regelgeving nodig is op het gebied van AI. Daarover is de data eenduidig, er is aparte wet- en regelgeving nodig. Echter is er ook hier sprake van een verschil, namelijk in de argumentatie waarom. R2 haakt aan op de definitie systeemtechnologie van de WRR:

“Nee, ik zit wel redelijk op de lijn van dat het systeemtechnologie is en dat betekent dat het ook niet zomaar een technologie is. Het is niet een soort van nieuwe radio die wordt uitgevonden of zo. Het is tegelijkertijd niet de allereerste systeem technologie die is ontwikkeld hè, dus ook daarin is er wel ervaring met hoe dat eerder is aangepakt.”

R3 kijkt meer naar het juridische aspect van het verhaal:

“Ja, kijk, Ik denk in een ideale wereld, is alles wat we al hebben afdoende, hè? Als je gewoon echt juridisch redeneert, dan is er natuurlijk al heel veel dat wat door AI systemen gebeurt, wat eigenlijk gewoon niet mag. In de sfeer van discriminatie of schade veroorzaken. Maar ja, we leven niet in de ideale wereld en ik denk dat in de wereld waarin we leven het wel nodig is om in ieder geval op dit moment wel dat soort regels te hebben. In ieder geval om het een beetje te kickstarten zeg maar en om gewoon te zorgen dat techbedrijven ook echt in actie komen.”

De verdiepende vragen over wat voor soort wet- en regelgeving er dan zou moeten zijn en wat daarvan de focus zou moeten zijn, worden eveneens verschillend beantwoord. R3 onderstreept dat het belang ervan in het regulerende effect zit, en zoals hierboven beschreven het de techbedrijven in actie kan laten komen:

“Ja, ik denk, het is niet een perfecte analogie hoor met rookverbod, ik zou dat er misschien niet per se voor gebruiken, maar heel even gewoon voor de conversatie om te laten zien dat je soms regelgeving niet per se maakt omdat je denkt dat dat helemaal strikt gezien juridisch de meest elegante oplossing is. Maar omdat die symboolwaarde, het feit dat je de regels maakt, op zichzelf soms ook al wel een regulerend effect kan hebben. Dus dat is vooral de analogie met rookverbod, denk ik.”

R2 bekijkt het meer vanuit verschillende fasen van de invoering van een AI-toepassing:

“Dat doen we nog afhankelijk van de tijdlijn van de ontwikkeling van AI. Als het ware zitten we nu eigenlijk niet in die fase waarin de eerste AI toepassingen echt een soort van in de samenleving terechtkomen op de markt worden gebracht. Wat er nu eigenlijk gebeurt, is dat we een soort van groot sociaal experiment met zijn allen aan het voeren zijn met AI. Wat eigenlijk natuurlijk een beetje van de zotte is. Als je gaat kijken naar bijvoorbeeld medicijnen, die moeten aan allerlei veiligheidseisen voldoen. Dus in het begin van het traject van de ontwikkeling van zo'n soort technologie zou ik zeggen dat je veel meer dus preventief moet nadenken. Op een later moment zouden dan ook juridische kaders helder gedefinieerd kunnen worden en ligt het voor de hand om mensen te gaan beboeten. Als de vraag is, waar zou je nu zelf de focus op leggen, dan zou dat toch van mij meer op het preventieve stuk zijn”

Dit sluit aan bij hetgeen hierover in D3 is opgeschreven. Daarin wordt namelijk gesteld dat het een samenspel van verschillende fasen in de tijd is, die van richtlijnen, heldere normen en standaarden uiteindelijk eindigt bij het toezicht. Hierbij spelen een impact-assessment en wettelijke waarborgen als de AVG² en de Grondwet ook een rol (Rathenau Instituut, 2022, pp. 8-10) Daaraan voegde R2 ook nog toe dat het een incrementeel proces is en het dus ook aan verandering onderhevig kan zijn.

² Algemene Verordening Gegevensbescherming

Samenvattend kan gesteld worden dat de respondenten vanuit verschillende argumentaties het nut en de noodzaak van wet- en regelgeving op het gebied van AI zien. Met andere woorden, het nut en de noodzaak van de governance van AI zien. De verschillen in opvattingen over en waarom en de relatie tussen de overheid en AI laten goed zien dat het thema AI leeft in verschillende disciplines, en daarmee dus ook impact heeft op (veel) verschillende disciplines. Hetgeen onderstreept wordt door D2. Tot slot zijn er verschillende dingen gesteld over de focus van de wet- en regelgeving, waarbij het regulerende effect en het preventieve deel expliciet benoemd zijn. Daarnaast is gesteld dat het een incrementeel proces is of dient te zijn, wat meerdere fases kent. Hetgeen overeenkomt met de stellingname in D3.

4.2 Functioneren en effectiviteit van adviesorganen

In deze paragraaf zullen de resultaten behandeld worden die gaan over het functioneren en de effectiviteit van adviesorganen. Allereerst werd de respondenten gevraagd wanneer zij achten dat een adviesorgaan goed functioneert. Daar kwamen verschillende aspecten mee naar boven, zo had R1 het over de diversiteit:

“Als de raad van diverse samenstelling is, dus niet alleen maar mensen uit één politieke hoek en niet alleen maar mensen uit de politiek-bestuurlijke hoek, maar ook mensen uit de wetenschap. dat geeft gewoon een mooie mix, zorgt ook wel voor nou ja, soms behoorlijke discussies. Maar ik denk, dat het resultaat vaak dan toch een breed gedragen resultaat is.”

Dit belang werd onderstreept door R2, die het tevens aanvulde met het belang van transdisciplinariteit:

“Ik wil daar een kleine toelichting op geven, zeg maar dat is dan wat inderdaad bedoeld wordt interdisciplinariteit, maar ook transdisciplinariteit wordt genoemd. Dus je wil interdisciplinariteit hebben, mensen met verschillende discipline achtergronden, dus ik wil een computerwetenschapper hebben, maar ook een filosoof bij wijze van spreken. En dan idealiter wil je ook iemand uit het bedrijfsleven hebben, iemand uit de publieke sector en iemand die voor maatschappelijke organisatie werkt, zodat je op die manier ook al die velden geborgd.”

Daarnaast benoemde R1 tevens het belang van een goede staf en bruikbare ingangen in Den Haag:

“Het is ook van belang dat er een goede staf is, want de raadsleden doen vaak voor 2 tot 4 dagen per week dit werk. De staf is dus erg belangrijk, niet alleen intern maar ook naar buiten. Zij hebben namelijk veel contact met bijvoorbeeld beleidsmedewerkers op de verschillende departementen, dus dat netwerk of die ingangen in Den Haag zijn ook belangrijk want beleidsmedewerkers zijn vaak met veel meer bezig dan alleen een advies van een orgaan. Dat geldt overigens ook voor raadsleden, dat het fijn is als zij goede ingangen in Den Haag hebben.”

R3 focuste meer op de opvolging van adviezen en de onafhankelijkheid van het orgaan:

“Ik zou toch wel zeggen dat het belangrijk is dat ze als een Autoriteit worden gezien en daardoor hun adviezen ook een bepaalde mate worden opgevolgd.”

“Misschien is dat dan gerelateerd aan een bepaalde mate van onafhankelijkheid, al wil ik daarmee niet zeggen dat zo een intern orgaan niet succesvol kan zijn, hoor,

maar er moet dan wel iets van onafhankelijk zijn om zo ook de integriteit van het advies op de een of andere manier te waarborgen.”

Tot slot kaartte R1 het belang van het streven naar een consensus over een bepaald advies aan:

“Het streven naar consensus, hè, dat is natuurlijk wel prettig als raadsleden niet eigen stokpaardjes blijven bereiden. Die hebben ze natuurlijk allemaal. Dat kan leiden tot moeilijke discussies of tot eventueel minderheidsstandpunten. En minderheidsstandpunten ja, daar schiet je helemaal niks mee op qua effectiviteit van je advisering. Dus dat moet ten allen tijde worden voorkomen.”

Daarnaast werd de respondenten gevraagd hoe zij kijken naar de functies van adviesorganen. Dat bracht vanuit de verschillende achtergronden van de respondenten interessante resultaten met zich mee. Alle respondenten waren het eens over de agenderende functie van adviesorganen:

“Nou toch dat je agenderend kan zijn, echt iets op de agenda zetten van de politiek, van het politieke bestuur, maar ook van de samenleving.” (R1)

“Juist omdat daar al die kennis samenkomt, hebben zij een overzicht en kunnen ze ook echt wel de regering en het Parlement bij wijze van spreken adviseren en daarbij dingen meegeven of onder de aandacht brengen.” (R2)

Daarbij gaf R3 de signalerende functie mee die volgens deze respondent hand in hand gaat met de agenderende functie:

“Ja, het woord dat in mij opkomt is wel signalerend, dus ook inderdaad wat grotere vragen oppakken, misschien wel voordat ze echt heel prangend worden. Dat is misschien ook eigenlijk wel gewoon een vorm van agenderen, maar in ieder geval die grotere vraagstukken signaleren en agenderen. Ik denk dat dat bij dit onderwerp wel, heel belangrijk is.”

Tot slot onderstreepte R1 wederom het belang van een goed netwerk, ditmaal in het licht van de agenderende functie:

“Maar hier speelt dan ook mee of je een goed netwerk hebt in Den Haag of niet. Kan je mensen daarmee zelf triggeren om iets op te pakken, bijvoorbeeld in een partijprogramma of komt het daardoor terug in een kamervraag?”

Andere woorden die gebruikt werden om de functie van adviesorganen te beschrijven waren prikkelend en adviserend. Daarbij onderstreepte R1 het belang van de mogelijkheid om ook ongevraagd te adviseren, wat tevens ten goede zou komen van de agenderende functie:

“Formeel zou een adviesorgaan dan zowel gevraagd als ongevraagd moeten kunnen adviseren. En niet alleen aan de regering, maar ook aan het parlement. En wat ik eerder zei van die adviezen, die komen vaker, ja, ook terecht bij andere overheden en andere organisaties, maatschappelijke organisaties soms ook, en die gaan daar weer hun eigen weg mee.”

Om samen te vatten, om hun agenderende, signalerende, prikkelende en uiteindelijk adviserende taak goed uit te kunnen voeren zijn een aantal zaken volgens de data van belang. Zo worden diversiteit en in het verlengde daarvan interdisciplinariteit en transdisciplinariteit genoemd, maar het belang van een goede staf en een goed netwerk werd tevens genoemd. Ook het waarborgen van onafhankelijkheid, het zijn van een autoriteit op een bepaald thema en het streven naar een consensus bij adviezen werden als belangrijk geïdentificeerd. Tot slot kaartte een respondent het belang van de mogelijkheid om ongevraagd te adviseren aan, in het licht van de agenderende functie.

4.3 Advisering op het gebied van AI

In deze paragraaf zullen de resultaten behandeld worden die gaan over advisering op het gebied van AI. Allereerst zal gekeken worden naar de inzichten die de respondenten en documenten hierover geven, waarna een blik zal worden geworpen op al bestaande manieren van advisering op het gebied van AI.

4.3.1 Inzichten uit de data

Allereerst is bij de respondenten nagegaan of zij de noodzaak van een adviserend orgaan op het gebied van AI al dan niet zien, en waarom zij dat al dan niet zien. R3 stelde dat het niet per se nodig is om een apart orgaan op te tuigen, maar dat het veel belangrijker is om er voor te zorgen er bij verschillende (adviserende) instanties expertise is op het gebied van AI:

“Ik was vorige week op een internationale conferentie van x, waar eigenlijk allerlei x instanties van de hele wereld waren. Er werd ook de vraag gesteld of we een aparte cyber x nodig hadden. Daar heb ik nee op gezegd, want volgens mij moet je die expertise, juist ook als x willen hebben en willen uitdragen en zo zie ik dat eigenlijk ook wel voor heel veel andere instanties en ook voor bijvoorbeeld toezichthouders om maar even een voorbeeld te noemen. Die expertise kun je gewoon niet alleen maar bij een orgaan beleggen. Dat moet juist op verschillende plekken, op die verschillende plekken moet die expertise en bewustzijn zijn.”

Echter nuanceerde deze respondent dat door de vergelijking te maken met de aparte wet- en regelgeving op het gebied van AI:

“Maar tegelijkertijd kan ik me ook een beetje, hetzelfde met die regelgeving, voorstellen dat er de komende jaren wel behoefte bestaat voor een orgaan dat daar misschien specifiek het voortouw in zou kunnen nemen.”

Ook R1 stelde dat het apart optuigen van een adviserend orgaan op het gebied van AI niet per se noodzakelijk zou zijn. Daarbij raakte deze respondent ook aan het adviesstelsel in Nederland, waar organen (vaak) gekoppeld zijn aan een departement:

“Nou, ik denk dat dit echt een overkoepelend thema is, dus ik zou het niet zo zeer onder willen brengen bij een apart adviesorgaan. Die adviesorganen zitten namelijk vaak aangehaakt aan een departement.”

Daar voegde deze respondent aan toe dat, binnen het huidige stelsel, dit onderwerp wel eventueel ergens anders zou kunnen worden ondergebracht. Zodat er wel een orgaan is dat in dat opzicht een coördinerende rol zal hebben:

"Maar ik zou zeggen zo'n onderwerp als AI, dat zie ik dan toch vooral belegd zijn bij de AWTI³, want die hebben al die invalshoek van technologie-innovatie en die hangen ook al boven bijna alle sectoren."

Tot slot nuanceerde deze respondent het antwoord door wederom de link te leggen met het adviesstelsel zoals dat in Nederland is vormgegeven door de Kaderwet Adviescolleges:

"Je zou dan zeggen dat voor zo'n onderwerp als AI ook een apart Ministerie van Digitale Zaken zou moeten komen, zodat die link misschien wat meer overkoepelend of explicieter gemaakt kan worden. Dat zeg ik puur persoonlijke titel, maar ik denk dat dat wel een goede zaak zou zijn. Overigens zou die kaderwet daar dan ook op toepassing moeten zijn, die regelt veel formele zaken."

Ook R2 stelde dat voor beide kanten van het verhaal wat te zeggen valt:

"Ik denk dat er voor alle twee de kanten natuurlijk iets te zeggen is. Aan de ene kant wil je sturing hebben of een soort van integraal beeld. Tegelijkertijd hebben we in het begin ook gehad over AI als systeemtechnologie. Het raakt aan allerlei verschillende soorten domeinen. Dat betekent ook dat je van al die domeinen ook vakinhoudelijk kennis moet hebben, om te kunnen begrijpen wat het effect van AI daarin gaat zijn. Dus je kan wel zeggen: laten we, bij wijze van spreken, een minister van AI in het leven roepen of zo, maar die heeft dan niet de vakinhoudelijke kennis van de andere ministeries."

Ook in D2 wordt de oproep voor een adviesorgaan gedaan, echter wordt er zelfs een stap verder gezet door te stellen dat er een beleidsinfrastructuur voor AI moet worden opgetuigd:

"Een beleidsinfrastructuur is in ieder geval nodig om de aankomende Europese AI-Verordening uit te voeren. Deze stelt dat de lidstaten één of meer nationale bevoegde autoriteiten aanwijzen om toezicht te houden op de toepassing en uitvoering van ai en als officieel contactpunt voor het publiek en andere actoren. Maar de opbouw van een beleidsinfrastructuur voor ai dient méér te omvatten, zo meent de WRR. Daarbij kan ons land zich laten inspireren door andere landen, die inmiddels AI-adviesraden en ambtelijke AI-bureaus instelden. De WRR acht het in deze fase te vroeg om een apart ministerie of een specifieke toezichthouder voor AI te bepleiten. Dit betekent echter niet dat de WRR de huidige status quo van het overheidsbeleid rondom AI als adequaat beoordeelt. Als eerstvolgende stap in de opbouw van een beleidsinfrastructuur bepleit de WRR daarom een coördinatiecentrum voor ai, dat aan beleidsdirecties, toezichthouders en uitvoeringsorganisaties een structuur biedt om regelmatig en rond uiteenlopende kwesties met elkaar in contact te treden en van elkaar te leren." (WRR, 2021, p. 20).

³ Adviesraad voor Wetenschap, Technologie en Innovatie

In navolging van de vraag over de noodzaak van en de ruimte voor een adviserend orgaan op het gebied van AI, zijn ook vragen over de functie, de focus van een dergelijk orgaan gesteld. Eveneens zijn de mogelijke inrichtingskeuzes van een dergelijk adviesorgaan behandeld, in de beantwoording van die vragen kunnen ook manieren voor de organisatie van kennismanagement en expertise geïdentificeerd worden, welke dus ook hier behandeld zullen worden. Allereerst de functie van een adviserend orgaan op het gebied van AI, daarin vinden R2 en R3 elkaar. Zij stellen beiden namelijk dat er wellicht een orgaan of instantie moet komen die vooral de onderlinge expertise op het gebied van AI (van bijv. andere adviesorganen) verbindt:

“Dus je wil eigenlijk denk ik naar een soort van tussenpositie, waar niet één iemand of een instelling verantwoordelijk wordt gemaakt voor alles wat met AI te maken heeft. Maar dat het meer een soort van integrerende rol heeft, om al die kennis die inderdaad op verschillende stukken wordt opgebouwd bij elkaar te integreren en te delen. Misschien is er dan bijna sprake van een soort van makelaarsfunctie.” (R2)

“Ik zou juist tegen een orgaan zijn waar je dan ook meteen alles van verwacht, op voorbeeld van AI en advies. Omdat je dan dus echt een hele organisatie moet gaan optuigen met een enorme staf. Ik denk ook dat dat nog steeds is wat ik zei, volgens mij is het belangrijk dat expertise hierover juist bij heel veel instellingen beter zijn weg gaat vinden. En dan kan zo'n commissie die bestaat uit allerlei mensen die die onderlinge kennis verbinden goed werken.” (R3)

Ook in D2 wordt de komst van een soort middle-man op het gebied van AI bepleit. Het wordt weliswaar geen adviesorgaan genoemd, maar een coördinatiecentrum:

“Het coördinatiecentrum kan kennis bij elkaar brengen, richting aanbrengen in de voor de overheid relevante vraagstukken, kansen en risico's rondom ai identificeren en een belangrijke coördinerende en faciliterende rol vervullen bij het opstellen van de bredere wetgevingsagenda die de WRR bepleit.” (WRR, 2021, p. 20).

Tevens suggereert dit stuk dat er een ministeriële onderraad moet komen, zodat er ook integrale afwegingen gemaakt kunnen worden:

“De WRR acht het daarom wezenlijk dat het centrum een politieke verankering kent, zodat er snel beleid kan worden gemaakt als dat nodig is en daartoe de politieke afstemming en sturing voorhanden zijn. Hiertoe beveelt de WRR de regering aan een ministeriële onderraad in te stellen waar zwaarwegende kwesties rondom digitalisering die om een integrale afstemming vragen, aan de orde komen.” (WRR, 2021, pp. 20-21)

Daar waar er dus een redelijke overeenkomst is over die verbindende rol, is R2 de enige respondent die zich expliciet uitsprekt over het beleggen van een een daadwerkelijk adviserende taak bij deze instantie:

“En die kun je ook als die kun je ook alsnog een adviesfunctie toeschrijven natuurlijk hè? Ik bedoel, juist omdat daar al die kennis samenkomt, hebben zij een overzicht en kunnen ze ook echt wel regering en Parlement bij wijze van spreken adviseren.”

R3 neemt daarover een wat meer genuanceerde stelling in. Deze respondent stelt namelijk dat er wel een rol kan liggen wat ongevraagde advisering betreft, maar dat wanneer er specifieke vragen zijn andere instellingen zich daar wellicht beter voor lenen:

“Dus niet per se op in opdracht werken. Je kan dan wel agenda opstellen en dat kan wel een beetje afgestemd worden met de regering en het parlement, maar in principe is het ook wel echt juist de kracht dat zij zelf dingen op de agenda zetten. Dat verlies je als je zegt: daar moet eigenlijk de hele overheid ook terecht kunnen voor allerlei adviesvragen, denk ik. Er zijn ook gewoon best wel specifieke adviezen nodig en dan lijkt het me helemaal niet gek als je daar dan weer juist een bepaalde instelling voor hebt, bijvoorbeeld het Rathenau instituut.”

Over de inhoudelijke focus van een dergelijk orgaan was R3 vrij uitgesproken. Deze respondent stelde dat het in ieder geval niet over juridische zaken moest adviseren, omdat daar al bestaande lijnen voor zijn:

“Dat op zich meer juridische advies, daar lopen natuurlijk al heel veel lijnen voor”

Dit werd later in het interview nog aangevuld met de stelling dat het niet te veel een niche gemaakt moet worden:

“AI en digitalisering, dat zit overal. Je moet daar dus ook weer niet te veel een niche gebied en een specialisatie van maken. Want je moet eigenlijk gewoon zorgen dat daarover over de hele linie heen expertise is.”

In de overige interviews is de inhoudelijke focus niet zo expliciet naar boven gekomen. Al kan in dit opzicht wel gesteld worden dat R1, R2 en D2 de consensus delen dat het een overkoepelend onderwerp of een systeemtechnologie is. Daaruit kan gedestilleerd worden dat een inhoudelijke afbakening volgens deze bronnen lastig te maken is.

Wat de mogelijke inrichtingskeuzes betreft zijn de respondenten relatief eensgezind. Zoals eerder al is gepresenteerd zijn de respondenten het het eens met elkaar dat er niet zo zeer een apart adviesorgaan dient te worden opgericht. Zij benoemen het idee van een meer coördinerend orgaan of commissie die de vraagstukken rondom AI coördineert:

“Die rol is dan meer coördinerend, van oh, je hebt dit probleem. Je moet daar eens kijken, zij weten daar wat van.” (R2)

R3 sluit daar nagenoeg naadloos bij aan, ter aanvulling geeft deze respondent het voorbeeld van de werkwijze van de Adviesraad Internationale Vraagstukken:

“Net zoals de AIV, dat is ongeveer zoals ook de Onderwijsraad is. Dus dat zijn dan inderdaad mensen die wel echt formeel benoemd zijn. En die zijn afkomstig uit wetenschap, het openbaar bestuur en kennisinstellingen. Volgens mij doet niemand die daar in zit dat ook voltijds. Dus het is juist de bedoeling dat ze naast hun werk en op persoonlijke titel, voor een periode van bijvoorbeeld 4 jaar en ze kunnen worden herbenoemd. Dat lijkt me wel een mooie vorm.”

Tevens past dit idee bij hetgeen in D2 wordt geopperd. Daarin beschrijft de WRR namelijk -zoals al eerder gepresenteerd- de noodzaak voor een coördinatiecentrum wat onder andere kennis bij elkaar kan brengen en richting kan aanbrengen in de voor de overheid relevante vraagstukken (WRR, 2021, pp. 20-21).

Eigenlijk is alleen R1 hier een andere mening toegedaan. Deze respondent stelt namelijk dat het onderwerp AI ondergebracht kan worden bij de Adviesraad voor Wetenschap, Technologie en Innovatie. Deze respondent is daarentegen wel de enige die een apart adviesorgaan bepleit, wanneer er een losstaand Ministerie van Digitale Zaken zou komen.

Samenvattend identificeren de bronnen een aantal belangrijke zaken, waarvan een centraal coördinerend orgaan het vaakst naar boven komt. Daarnaast laten de bronnen zien dat het onderwerp AI dermate veel domeinen raakt, dat een sterke inhoudelijke afbakening niet goed te maken is. Dit maakt het volgens de bronnen van belang dat er binnen die verschillende domeinen en instanties zelf kennis en expertise is op het gebied van AI. Tot slot is er wel consensus binnen de verzamelde data over het belang van een dergelijk coördinerend -dan wel adviserend- orgaan op het gebied van AI. Dit wordt wederom vanuit verschillende perspectieven onderbouwd.

4.3.2 Bestaande manieren van advisering op het gebied van AI

Op het moment zijn er al landen met officiële organen die zich bezighouden met de advisering op het gebied van AI, zo zijn Spanje en de VS internationale koplopers op dit gebied. Daarnaast zijn er ook voorbeelden van landen waarbij de advisering meer versnipperd is, zoals Nederland. In dit deel van het resultatenhoofdstuk zal vooral gekeken worden naar de verschillende keuzes die gemaakt zijn ten aanzien van het organiseren van (expert)advisering op het gebied van AI.

Allereerst de VS, bij monde van een executive order heeft president Joe Biden het Department of Homeland Security opgedragen om een orgaan in het leven te roepen. Dit orgaan draagt de naam *The Artificial Intelligence Safety and Security Board*. De naam suggereert dat de inhoudelijke afbakening van dit adviesorgaan veel te maken heeft met de veiligheid van de VS. Het presidentieel decreet onderstreept dat. Daarin wordt gesteld dat het doel van dit orgaan het voorzien in advies, informatie en aanbevelingen betreffende het verbeteren van de veiligheid, weerbaarheid en de reactie op incidenten gerelateerd aan AI-gebruik in de kritieke infrastructuur. De inhoudelijke focus wordt dus expliciet aangebracht op de veiligheid, weerbaarheid en reactie op incidenten in de kritieke infrastructuur. Daarmee is ook het publiek van het advies redelijk smal. Er wordt gesteld dat er geadviseerd wordt aan de Minister van Binnenlandse Veiligheid en aan de federale community rondom de kritieke infrastructuur. Tot slot de gemaakte inrichtingskeuzes. Het decreet schrijft voor dat de leden van het orgaan AI-experts uit de private sector, academische sector en publieke sector/het bestuur dienen te komen. Kortom, AI-experts uit verschillende werkvelden (The White House, 2023).

Het andere voorbeeld, Spanje, dat kiest een andere insteek dan de VS. De AESIA, het Spaanse Agentschap voor de Supervisie van Kunstmatige Intelligentie, is een orgaan dat opereert binnen het departement van Digitale Transformatie en is dus geen onafhankelijk orgaan. Het doel van dit orgaan is het uitvoeren van supervisie (op) en advisering, het creëren van bewustzijn en het geven van trainingen aan publiek- en privaatrechtelijke entiteiten. Dit

allemaal in het kader van de implementatie van alle nationale en EU wet- en regelgeving op het gebied van AI en algoritmen. Daarbij heeft dit agentschap ook de bevoegdheid om inspecties uit te voeren, zaken te verifiëren en om te sanctioneren specifiek op het gebied van AI. Dit orgaan heeft dus als doel om vanuit de overheid andere (overheids-)instellingen te adviseren op de genoemde thema's, en kan naleving afdwingen met de genoemde bevoegdheden. Dit agentschap wordt niet gevuld met (externe) expertise, de betreffende wet schrijft voor dat het gaat om ambtenaren, met bijbehorende hiërarchie. Aan het hoofd staat dus de politiek verantwoordelijke voor het Ministerie van Digitale Transformatie. Wel bestaat de mogelijkheid om voor bepaalde beleidsterreinen (bijvoorbeeld defensie of de zorg) externe experts te betrekken bij de uitvoering van hun taken (Real Decreto 729/2023, Por El Que Se Aprueba El Estatuto de La Agencia Española de Supervisión de Inteligencia Artificial., 2023)

Tot slot de Nederlandse situatie, daarin valt op dat een centraal orgaan ontbreekt. De advisering is versnipperd over allerlei instanties. Goede voorbeelden hiervan zijn D2 en D3, een adviesrapport en een handreiking van respectievelijk de Wetenschappelijke Raad voor het Regeringsbeleid en het Rathenau Instituut. Tevens verschilt het publiek van deze twee documenten, de eerste is gericht aan de Nederlandse regering en bevat concrete adviezen daaraan. Het stuk van het Rathenau Instituut is gericht aan de Eerste Kamer en dient als een notitie ter ondersteuning van de werkgroep AI. Deze versnippering is tekenend en wordt andermaal aangetoond door de adviesaanvraag *AI, de toekomst van werk en sociaaleconomische implicaties*. Dit is een adviesaanvraag van het Nederlandse kabinet aan de Sociaal Economische Raad, een gerichte vraag aan een apart adviesorgaan (Ministerie van Binnenlandse Zaken en Koninkrijksrelaties, 2024).

Er zijn dus verschillende manieren hoe verschillende landen dit aanvliegen. Van organen die gevuld zijn met externe expertise, tot een versnipperd adviesaanbod. Ook organen die onderdeel zijn van een ministerie komen voorbij. Dit is tekenend voor de situatie waarin de wereld verkeert omtrent AI, het is een relatief nieuwe technologie en dus is er (nog) geen blauwdruk voor hoe het (expert)adviseringsvraagstuk daaromtrent benaderd kan worden.

4.4 Reflectie op de empirische verwachtingen

In deze paragraaf zal de verzamelde data naast de empirische verwachtingen die voortvloeien uit het theoretisch kader worden gelegd, met als doel de verwachtingen te ontcrachten dan wel te bevestigen. Dit zal het onderliggende fundament zijn voor het volgende hoofdstuk, de conclusie en discussie.

De eerste empirische verwachting betreft de vraag of er wel aparte wet- en regelgeving nodig is op het gebied van AI, en in het verlengde daarvan of er aparte advisering nodig zou zijn. De verwachting is wat deze vraag betreft uitgekomen. Het belang van aparte wet- en regelgeving werd onderschreven, en tevens is het belang van aparte advisering op het gebied van AI benoemd. Dit is in lijn met hetgeen Chauhan (2022) daarover op papier heeft gezet. Ten tweede is hier de verwachting uitgesproken dat het proces van de totstandkoming van de wet- en regelgeving over AI centraal zal staan en dat een adviserend orgaan daarbij vanuit

verschillende visies een rol kan spelen. Deze verwachting is deels ontkracht en deels bevestigd. Het proces van de totstandkoming rondom de wet- en regelgeving is niet als centraal belang naar boven gekomen uit de data, hetgeen wel door onder andere. Sturm, Weyerer en Wirtz (2020) is gepropageerd. De data bevestigt wel het belang van de verschillende visies, en komt daarmee onder andere overeen met Almeida en Gasser (2017), Sturm, Weyerer en Wirtz (2020) en Choi en Park (2023). Ook de verwachting dat het orgaan overkoepelend dient te zijn wordt door de data bevestigd, en is daarmee in overeenstemming met Sturm, Weyerer en Wirtz (2020). Tot slot de factoren die als belangrijk worden gezien op het gebied van advisering over AI. De verwachting dat in ieder geval een diversiteit aan kennis zal worden aangedragen gezien het feit dat AI een systeemtechnologie is en veel vlakken raakt, zoals uit het theoretisch kader is gebleken, is daar uitgesproken. Deze verwachting komt bijna helemaal overeen met wat uit de verzamelde data is gebleken. De data noemde echter nog meer factoren, zoals de transdisciplinariteit. Tevens liet de data zien dat een soort expertise, de juridische, niet nodig zou zijn omdat daar al genoeg mogelijkheden betreffende advisering voor zijn.

De tweede set empirische verwachtingen die zijn uitgesproken, slaan op het inrichtingsvraagstuk. Daar is een drietal mogelijkheden (geen nieuw orgaan, wel nieuw orgaan, taak bij een bestaand orgaan onderbrengen) beschreven. De verwachting dat de data zal wijzen op een van de laatste twee mogelijkheden is daar gesteld. Toevoeging daaraan is dat de doorslaggevende factor voor de keuze tussen die twee, de beoogde rol van het orgaan zal zijn. Hierover is de data conflicterend, een respondent zegt dat het ondergebracht kan worden bij een bestaand orgaan. Daar waar twee andere respondenten het oprichten van een coördinerende instantie bepleiten. Daarbij komt ook nog dat een van de documenten wel oppert om een aparte adviesraad in te stellen. Deze verwachting is dus niet in lijn met hetgeen uit de data is gebleken. Dit is ook het geval voor de verwachting aangaande het belang van de benoeming van de leden. Hetgeen daarover uit de theorie is gebleken, veronderstelt dat dit een relatie heeft met de autoriteit van de raad. De data ondersteunt deze claim niet, echter onderschrijft zij wel het belang van diezelfde autoriteit. Aangevuld met diversiteit en onafhankelijkheid van het orgaan. Tot slot is de verwachting uitgesproken dat het kennismanagement enige mate van incrementaliteit dient te hebben, omdat de techniek onderhevig is aan verandering en daarmee periodieke bijsturing gewenst is. De data bevestigt het belang van een niet in beton gegoten manier van kennismanagement, zoals bijvoorbeeld een vaste samenstelling. Daarmee is het in lijn met hetgeen Osborn (2004) hierover stelt.

5. Conclusie & Discussie

Dit hoofdstuk is opgebouwd uit twee delen. Allereerst zal een antwoord op de hoofdvraag geformuleerd worden. Daarna zullen in het kader van de discussie de tekortkomingen van dit onderzoek en aanbevelingen voor vervolgonderzoek behandeld worden.

5.1 Conclusie

In deze paragraaf zal een antwoord op de hoofdvraag geformuleerd worden. Dit antwoord vloeit voort uit de vergaarde data die naast de theorie is gelegd. Daarin is getracht verschillende belangrijke factoren voor advisering op het gebied van AI te identificeren, alsmede de mogelijke inrichtingsvormen voor een adviserend orgaan op het gebied van AI te identificeren. Tevens is een blik geworpen op hoe advisering op het gebied van AI in de VS, Spanje en Nederland eruit ziet. De hoofdvraag van dit onderzoek luidt: *Wat zijn de mogelijkheden voor het inrichten van een adviserend orgaan op het gebied van AI?*

In het antwoord op deze vraag valt op basis van de literatuurstudie en de empirische bevindingen grofweg een onderscheid te maken tussen twee mogelijkheden waarbij de geïdentificeerde belangrijke factoren gewaarborgd zijn. De eerste mogelijkheid is het beleggen van de adviserende taak op het gebied van AI bij een al bestaand adviesorgaan. In de Nederlandse context werd daarvoor de Adviesraad voor Wetenschap, Technologie en Innovatie genoemd. Generaliserend kan er dus geconcludeerd worden dat wanneer men voor deze mogelijkheid kiest, het onderwerp AI ondergebracht dient te worden bij een orgaan dat hoofdzakelijk met deze thema's bezig is.

De tweede mogelijkheid gaat niet zo zeer uit van een orgaan dat alleen adviseert, maar wat meer een makelaarsfunctie heeft. Met andere woorden, het brengt vraag en aanbod samen. Daarbij zou wel ruimte moeten zijn voor een adviserende taak, en het liefst is dat dan ongevraagd. Door op deze manier vorm te geven aan de inrichting van een adviserend orgaan op het gebied van AI, is er ruimte voor de experts om zelf te adviseren op bijvoorbeeld de richting van regeringsbeleid omtrent AI, maar tevens om specifieke adviesvragen op het snijvlak van AI en een ander domein door te zetten naar de organisaties die daarover kennis en expertise hebben. Ook op dit orgaan kan de Kaderwet Adviescolleges van toepassing zijn, zodat de formele zaken omtrent bijvoorbeeld benoemingen en onafhankelijkheid. Tot slot is het van belang dat er sprake is van interdisciplinariteit en transdisciplinariteit, zodat er als het ware een intellectuele kopgroep ontstaat op het gebied van AI.

Beide van deze mogelijkheden komen niet overeen met hetgeen de VS, Spanje of Nederland op dit moment doen op het gebied van advisering over AI. Dat betekent concreet dat er of geen sprake is van het delen van best practices, of dat de goudader wat advisering op het gebied van AI betreft nog niet gevonden is.

5.2 Tekortkomingen onderzoek

In dit onderzoek kunnen een aantal tekortkomingen gesignaleerd worden, die maken dat de conclusie en de reflectie op de empirische verwachtingen niet zo hard gemaakt kunnen worden. Allereerst de diversiteit in en de hoeveelheid respondenten. Er is getracht een zo divers mogelijk palet aan respondenten te creëren, zodat tal van expertises geraadpleegd konden worden. De realiteit is echter dat dat niet gelukt is, dit terwijl er negentien unieke experts benaderd zijn voor deelname aan dit onderzoek. Zo had het bijvoorbeeld goed geweest als een bestuurskundige die kijkt naar de relatie tussen de overheid en adviesorganen, en daarmee naar de effectiviteit van adviesorganen, had deelgenomen.

Tevens hanteert dit onderzoek een brede hoofdvraag, dat is in dit geval inherent aan het verkennende karakter van het onderzoek. Dit kan er echter voor zorgen dat er niet diep genoeg is ingegaan op bepaalde aspecten in zowel de literatuurstudie als de dataverzameling. Het gevolg hiervan is dat mogelijk niet alle belangrijke factoren behandeld zijn of even goed aan bod komen. Dit in combinatie met het geringe aantal respondenten maakt dat ondanks de brede insteek van de hoofdvraag, ineens op een klein aantal mogelijkheden de diepte in wordt gegaan. Hetgeen niet bevorderlijk is voor de houdbaarheid van de generalisatie van de conclusie.

Tot slot is het inherent aan kwalitatief onderzoek dat de herhaalbaarheid lastiger te waarborgen is dan bij kwantitatief onderzoek. Dat maakt dat de validiteit en betrouwbaarheid van dit onderzoek in het gedrang kunnen komen. Er is getracht dit te ondervangen door de respondenten en documenten te voorzien van een objectieve beschrijving, maar als de context onderhevig is geweest aan veranderingen, heeft dat ook niet heel veel nut meer.

5.3 Aanbevelingen vervolgonderzoek

Allereerst is aan te bevelen om onderzoek te doen naar de effectiviteit van verschillende soorten adviesorganen. Dit zodat verkennende onderzoeken als deze makkelijker uit te voeren zijn. Dat is niet alleen ten behoeve van onderzoekers, maar ook van de samenleving omdat AI niet de eerste, noch de laatste systeemtechnologie zal zijn waar de overheid wat mee moet.

Ten tweede is het aan te bevelen om onderzoek te doen naar welk bestaand adviesorgaan zich het best leent om het overkoepelende thema AI op zich te krijgen. Daarnaast is het aan te bevelen om het creëren van een orgaan dat als makelaar en adviseur optreedt nader te onderzoeken. Dit zijn de twee voornaamste conclusies uit deze verkenning, waarbij het zwaartepunt op het tweede ligt. Tevens is deze tweede optie interessant omdat er nog geen orgaan is waarin deze twee functies verenigd worden. Tot slot zou het ook interessant zijn om bij dat onderzoek te betrekken hoe gewaarborgd wordt dat de kennis en expertise over AI in verschillende domeinen gewaarborgd blijft. Dit zodat de makelaarsfunctie niet in het gedrang kan komen.

Literatuurlijst

1. Aarts, J., Gerth, E., Ludwig, D., Maat, H., & Macnaghten, P. (2022). The Dutch see Red: (in)formal science advisory bodies during the COVID-19 pandemic. *Humanities And Social Sciences Communications*, 9(1). <https://doi.org/10.1057/s41599-022-01478-w>
2. Acosta, M., Nestore, M., Jarquín-Solís, M. E., & Doubleday, R. (2022). A typology of advisory bodies in legislatures and research perspectives. *The Journal Of Legislative Studies*, 1–26. <https://doi.org/10.1080/13572334.2022.2070985>
3. Almeida, V. A. F., Mendes, L. S., & Doneda, D. (2023). On the Development of AI Governance Frameworks. *IEEE Internet Computing*, 27(1), 70–74. <https://doi.org/10.1109/mic.2022.3186030>
4. Andrews, L. (2018). Public administration, public leadership and the construction of public value in the age of the algorithm and ‘big data’. *Public Administration*, 97(2), 296–310. <https://doi.org/10.1111/padm.12534>
5. Barth, T. J., & Arnold, E. (1999). Artificial Intelligence and Administrative Discretion. *The American Review Of Public Administration*, 29(4), 332–351. <https://doi.org/10.1177/02750749922064463>
6. Birkstedt, T., Minkkinen, M., Tandon, A., & Mäntymäki, M. (2023). AI governance: themes, knowledge gaps and future agendas. *Internet Research*, 33(7), 133–167. <https://doi.org/10.1108/intr-01-2022-0042>
7. BOE-A-2023-18911 Real Decreto 729/2023, de 22 de agosto, por el que se aprueba el Estatuto de la Agencia Española de Supervisión de Inteligencia Artificial. (z.d.). https://boe.es/diario_boe/txt.php?id=BOE-A-2023-18911
8. Boeije, H., & Bleijenbergh, I. (2019). *Analyseren in kwalitatief onderzoek* (3de editie). Boom Lemma.
9. Boeije, H. (2014). *Analyseren in kwalitatief onderzoek*. Amsterdam: Boom
10. Bryman, A. (2016). *Social Research Methods*, Oxford University Press, Oxford. (5th edition).
11. Chauhan, K. D. S. (2022). From ‘What’ and ‘Why’ to ‘How’: An Imperative Driven Approach to Mechanics of AI Regulation. *Global Jurist*, 23(2), 99–124. <https://doi.org/10.1515/gj-2022-0053>
12. Choi, H., & Park, M. J. (2023). To govern or be governed: an integrated framework for AI governance in the public sector. *Science And Public Policy*, 50(6), 1059–1072. <https://doi.org/10.1093/scipol/scad045>
13. Crowley, K. Head, B. Expert advisory councils in the policy system. In Brans, M. Geva-May, I. and Howlett, M (eds) 2017 *Routledge Handbook of Comparative Policy Analysis*, NY and London: Routledge, pp. 181-198.
14. DPG Media, De Volkskrant (z.d.). <https://www.volkskrant.nl/nieuws-achtergrond/belastingdienst-gebruikte-algoritme-dat-lage-inkomens-selecteerde-voor-extra-fraudecontroles~bac84336/>
15. Gasser, U., & Almeida, V. A. F. (2017). A Layered Model for AI Governance. *IEEE Internet Computing*, 21(6), 58–62. <https://doi.org/10.1109/mic.2017.4180835>
16. Giest, S. N., & Klievink, A. J. (2022). More than a digital system: how AI is changing the role of bureaucrats in different organizational contexts. *Public Management Review*. <https://doi.org/10.1080/14719037.2022.2095001>
17. Henman, P. (2020) Improving public services using artificial intelligence: possibilities, pitfalls, governance, *Asia Pacific Journal of Public Administration*, 42:4, 209-221, <https://doi.org/10.1080/23276665.2020.1816188>

18. Jacobs, M., & Simon, J. (2022). Assigning Obligations in AI Regulation: A Discussion of Two Frameworks Proposed By the European Commission. *Digital Society*, 1(1). <https://doi.org/10.1007/s44206-022-00009-z>
19. Kuziemski, M., & Misuraca, G. (2020). AI governance in the public sector: Three tales from the frontiers of automated decision-making in democratic settings. *Telecommunications Policy*, 44(6), 101976. <https://doi.org/10.1016/j.telpol.2020.101976>
20. Ministerie van Binnenlandse Zaken en Koninkrijksrelaties. (2024, 18 januari). SER-adviesaanvraag 'AI, de toekomst van werk en sociaaleconomische implicaties'. Rapport <https://www.rijksoverheid.nl/documenten/rapporten/2024/01/18/adviesaanvraag-ser-ai-algoritmes-en-data-de-toekomst-van-werk-en-sociaaleconomische-implicaties>
21. Nordström, M. (2021). AI under great uncertainty: implications and decision strategies for public policy. *AI & SOCIETY*, 37(4), 1703–1714. <https://doi.org/10.1007/s00146-021-01263-4>
22. NOS. (2021a, 15 januari). Kabinet-Rutte III gevallen; Wiebes helemaal weg. NOS. <https://nos.nl/collectie/13855/artikel/2364513-kabinet-rutte-iii-gevallen-wiebes-helemaal-weg>
23. NOS. (2021b, 18 januari). Toeslagenaffaire. De ellende uitgelegd. [Video]. NOS. <https://nos.nl/op3/collectie/13855/video/2364958-toeslagenaffaire-de-ellende-uitgelegd>
24. Osborne, M. (2004). Making the most of the best: Formulating a knowledge management strategy for a government advisory body. *Australian Journal Of Public Administration*, 63(3), 43–52. <https://doi.org/10.1111/j.1467-8500.2004.00389.x>
25. Parlementaire ondervragingscommissie Kinderopvangtoeslag. (2020, 17 december). Ongekend onrecht. Geraadpleegd van https://www.tweedekamer.nl/kamerleden_en_commissies/commissies/pok
26. Rathenau Instituut. (2020). Grip op algoritmische besluitvorming bij de overheid. In *Notitie Ter Ondersteuning van de Werkgroep AI*. https://www.rathenau.nl/sites/default/files/2021-05/Notitie_Eerste_Kamer_Grip_op_algoritmische_besluitvorming_overheid_Rathenau_Instituut.pdf
27. RTL Nieuws. (2020, 17 december). Hoe kon het zo misgaan bij de toeslagenaffaire? Volg de presentatie live. <https://www.rtlnieuws.nl/nieuws/politiek/artikel/5203740/presentatie-vernietigend-rapport-toeslagenaffaire>
28. RTL Nieuws. (2022, 15 september). College voor de Rechten van de Mens: Belastingdienst discrimineerde mensen van buitenlandse afkomst. <https://www.rtlnieuws.nl/nieuws/politiek/artikel/5333439/college-voor-de-rechten-van-de-mens-belastingdienst-discrimineerde>
29. Sharma, S. (2023). Trustworthy Artificial Intelligence: Design of AI Governance Framework. *Strategic Analysis*, 47(5), 443–464. <https://doi.org/10.1080/09700161.2023.2288994>
30. Sigfrids, A., Nieminen, M., Leikas, J., & Pikkuaho, P. (2022). How Should Public Administrations Foster the Ethical Development and Use of Artificial Intelligence? A Review of Proposals for Developing Governance of AI. *Frontiers in Human Dynamics*, 4. <https://doi.org/10.3389/fhumd.2022.858108>
31. Van Rij, M.L.A. (2022, 30 mei). Nadere verzoeken Fraude signalering Voorziening [kamerbrief]. Geraadpleegd van <https://www.rijksoverheid.nl/documenten/kamerstukken/2022/05/30/kamerbrief-reactie-op-verzoeken-over-fraudesignaleringsvoorziening>

32. Bernd W. Wirtz, Jan C. Weyerer & Benjamin J. Sturm (2020) The Dark Sides of Artificial Intelligence: An Integrated AI Governance Framework for Public Administration, *International Journal of Public Administration*, 43:9, 818-829, <https://doi.org/10.1080/01900692.2020.1749851>
33. The White House (2023, 30 oktober). Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence. The White House. <https://www.whitehouse.gov/briefing-room/presidential-actions/2023/10/30/executive-order-on-the-safe-secure-and-trustworthy-development-and-use-of-artificial-intelligence/>
34. WRR., oa. Prins, C. (2021). Opgave AI: de nieuwe systeemtechnologie.
35. Zouridis, S., Van Eck, M., & Bovens, M. (2019). Automated discretion. *Social Science Research Network*. <https://doi.org/10.2139/ssrn.3453068>

Bijlagen

1. Tabel: Schematische weergave kenmerken Acosta et al (2022)

<i>Kenmerk</i>	<i>Mogelijkheden</i>
Het politieke systeem	<ul style="list-style-type: none"> - Presidentieel - Parlementair
De inrichting van het kamerstelsel	<ul style="list-style-type: none"> - Eenkamer - Tweekamer
Het mandaat	<ul style="list-style-type: none"> - Enkelvoudig (alleen adviserend) - Meervoudig (meer dan alleen adviserend, bijv. voorschrijvend of adviserend voor meerdere groepen)
Het bedoelde publiek	<ul style="list-style-type: none"> - De wetgevende macht - De samenleving - Academici en studenten - De uitvoerende macht
Het doel van het advies	<ul style="list-style-type: none"> - Beschrijvend (sec adviserend) - Voorschrijvend (beleidswijzigingen uitwerken en voorschrijven) - Gemixt
De samenstelling van het orgaan	<ul style="list-style-type: none"> - Alleen wetenschappers - Alleen ambtenaren - Alleen leden van de wetgevende macht - Wetgevers en wetenschappers - Ambtenaren en wetenschappers - Wetenschappers, ambtenaren en wetgevers
De benoemingsprocedure	<ul style="list-style-type: none"> - Politieke benoeming - A-politieke benoeming

Aangepast en vertaald overgenomen uit Acosta et al (2022, pp. 6-9).

2. Topiclijsten & interviewprotocol

2.1 Interviewprotocol

Het interviewprotocol slaat op de formele zaken van het vergaren van data op deze manier. Naast het netjes voorstellen en het benoemen van de structuur en opbouw van het interview, worden hiermee de rechten van de respondent kenbaar gemaakt. Tevens wordt gevraagd om de toestemming van de respondent om het interview op te nemen en de data te verwerken. In het kader daarvan zijn de respondenten van het volgende op de hoogte gesteld:

- Het interview zal worden opgenomen, waarna het getranscribeerd zal worden zodat de data kan worden gebruikt in het onderzoek;
- Het transcript zal alleen (wanneer nodig) gedeeld worden met de begeleidend hoogleraar;
- Wanneer het onderzoek is beoordeeld zullen de transcripten en opnames worden verwijderd;
- De data zal geanonimiseerd worden alvorens het wordt gebruikt in het onderzoek, zodoende is de data niet te herleiden tot de respondent;
- De respondent heeft gedurende het interview alsmede gedurende de rest van het onderzoek het recht om alsnog af te zien van deelname, zonder opgave van reden;
- Mocht de respondent naderhand afzien van deelname, dan zal de data worden verwijderd en zodoende niet gebruikt in het onderzoek;
- Mocht de respondent vragen of opmerkingen hebben, dan is de onderzoeker bereikbaar, tevens is het voor de respondent mogelijk om desgewenst contact op te nemen met de begeleidend hoogleraar;
- Geeft u toestemming voor opname van dit interview en verwerking van de data in het onderzoek?

2.2 Topiclijst experts op het gebied van de governance van AI

- Relatie tussen AI en de overheid?
- Wet- en regelgeving voor AI, apart voor opstellen of is huidige wet- en regelgeving op het gebied van technologie en innovatie genoeg?
 - Waarom wel of niet?
- Wat voor soort wet- en regelgeving is goed voor AI? (bijv. sanctionerend of voorschrijvend)
 - Waarom?
- Kan een adviserend orgaan op het gebied van AI een aanvulling zijn hierop?
 - Waarom wel of niet?
- Wat zou de focus van wet- en regelgeving op het gebied van AI moeten zijn? (bijv. meer het technisch fundament of de maatschappelijke impact)
- Hypothetische situatie: er is een adviserend orgaan op het gebied van AI. Wat zou de inhoudelijke focus van dit orgaan moeten zijn (bijv. gevraagd advies op wetsvoorstellen over AI, of ongevraagd advies op de maatschappelijke impact van een AI-toepassing etc).

- Wat zou de functie van dit orgaan moeten zijn? (bijv. agenderend, voorschrijvend of puur het samenbrengen van verschillende expertises)

2.3 Topiclijst experts op het gebied van het adviesstelsel en adviesorganen

- Wanneer functioneert een adviesorgaan goed?
- Wat maakt dat een adviesorgaan goed functioneert?
 - Welke factoren spelen daarbij een rol?
- Wat is de visie op de functie van adviesorganen?
 - Valt dat te verenigen met een onderwerp als AI?
- Wat is de visie op het adviesstelsel?
 - Waarom zou een orgaan op het gebied van AI daar al dan niet tussen passen?
- Wanneer men besluit tot het oprichten van een adviesorgaan op het gebied van AI, wat zijn dan factoren waarmee men rekening moet houden bij de inrichting?
- Wat zijn inrichtingskeuzes die gemaakt kunnen worden?
 - Op welke manier zouden die gemaakt kunnen/moeten worden wanneer men een adviserend orgaan op het gebied van AI wil oprichten?
- Wat zijn andere factoren die invloed hebben op de vorm van een adviesorgaan, die mogelijk ook hierop van toepassing zijn?

3. Specificatie respondenten

Respondentnummer	Expertisegebied
R1	Adviesstelsel & adviesorganen
R2	Advisering & beleid op het gebied van technologie en innovatie (oa. AI)
R3	Governance van AI

4. Specificatie documenten

Titel	Documentnummer	Auteur(s)	Type document	Waar gevonden
Kaderwet adviescolleges	D1	Ministeries van BZK ⁴ en J&V ⁵	Wetgeving	Aangedragen door een respondent
Opgave AI. De nieuwe systeemtechnologie	D2	WRR ⁶	Adviesrapport	Aangedragen door een respondent & bekend van een eerder gevolgd vak
Grip op algoritmische besluitvorming. De rol van de Eerste Kamer	D3	Rathenau Instituut	Adviesrapport	Bekend van eerdere werkzaamheden in de Eerste Kamer
Executive Order on the Safe, Secure, and Trustworthy Development and Use of Artificial Intelligence	D4	The Joseph R. Biden jr. administration (DHS ⁷)	Presidentieel decreet (executive order)	Google zoekopdracht "AI-adviesraad VS"
Statute of the Spanish Agency for the Supervision of Artificial Intelligence.	D5	Het Spaanse ministerie van Digitale Transformatie	Wetgeving tbv het instellen van het orgaan	Google zoekopdracht "AI-adviesraad Spanje"

⁴ Binnenlandse Zaken en Koninkrijksrelaties

⁵ Justitie & Veiligheid

⁶ Wetenschappelijke Raad voor het Regeringsbeleid

⁷ Department of Homeland Security