

Op de rit met algoritmen?

Een empirisch onderzoek naar percepties van legitimiteit en de doorwerking van het Wmo-voorspelmodel in beleidsvorming

Student: Roos den Hollander

Studentnummer: 582672

E-mail: 582672rh@eur.nl

Begeleider: dr. ir. Jasper Eshuis

Tweede beoordelaar: dr. Rebecca Moody

Stageorganisatie: ScienceWorks

Master Bestuurskunde Beleid & Politiek

Erasmus Universiteit Rotterdam

Aantal woorden: 17.519

Datum: 28-07-2022

Voorwoord

Graag neem ik u mee in de totstandkoming van deze thesis over algoritmen in het openbaar bestuur en percepties van legitimiteit. Als er iets is wat deze thesis mogelijk heeft gemaakt, dan is het doorzettingsvermogen. Het aantal keren dat ik met een diepe en vermoeide zucht mijn laptop afsloot na een dag hard werken, is niet op twee handen te tellen. Maar het is waar wat men zegt: zonder wrijving geen glans.

Toen ik aan de master Bestuurskunde begon, stond het onderwerp algoritmen niet bovenaan mijn lijst met interesses. Echter, eenmaal ondergedompeld in het onderwerp groeide de fascinatie al snel. Wat daarbij heeft geholpen is mijn stage bij ScienceWorks waar ik niet alleen meehielp bij het organiseren van een Leergang over algoritmen binnen de overheid, maar ook eraan mocht deelnemen. Daar ben ik erg dankbaar voor!

Ook de Vereniging van Nederlandse Gemeenten (VNG) ben ik dankbaar. In de zoektocht naar een casus voor het onderzoek, kwam ik de landelijke lancering van het Wmo-voorspelmodel tegen. Ik ben erg dankbaar dat de VNG open stond voor deze samenwerking. Specifiek wil ik Janneke Lummen bedanken voor haar tijd en inzet, voor het benaderen van respondenten, maar ook voor de vrijheid om te doen wat ik voor ogen had. Ook wil ik mijn bijzondere dankbaarheid uiten naar de respondenten die wilden meewerken aan dit onderzoek.

Niet allerminst wil ik mijn scriptiebegeleider Jasper Eshuis bedanken. Ondanks dat ik soms met meer vragen naar huis ging na afloop van scriptiekringen dan ik vooraf had, heb ik altijd de feedback en scriptiekringen als zeer prettig ervaren. Het zette me verder aan het denken tot frustratie aan toe, maar ook dat hoort bij die wrijving. Daarbij wil ik mijn medestudenten bedanken die mij ook veel feedback hebben gegeven en waarmee ik al die afstudeerstress kon delen. Dat doorzettingsvermogen wat ik eerder noemde, geldt overigens niet alleen voor mij, maar ook zeker voor hen.

Ten slotte wil ik mijn familie en vrienden bedanken die ook mijn frustraties gevraagd en ongevraagd aanhoorden, maar die tevens voor de nodige afleiding zorgden. Zo hebben alle bovengenoemden ervoor gezorgd dat ik met trots terugkijk op het afgelopen jaar met deze thesis en hopelijk een diploma als resultaat.

Dit gezegd hebbende wens ik u veel leesplezier toe.

Roos den Hollander

Zevenbergen, 22 juli 2022

Samenvatting

In het openbaar bestuur is een digitale transformatie gaande. Door het toegenomen vermogen om enorme hoeveelheden data te verzamelen, produceren, verwerken en combineren, zijn onder andere complexere scenarioberekeningen mogelijk. Zo maakt de overheid steeds meer gebruik van algoritmen binnen het beleidsproces, bijvoorbeeld om het te vormen beleid te onderbouwen. Dit maakt het openbaar bestuur beter in staat te anticiperen op toekomstige ontwikkelingen. In de praktijk blijkt echter sprake te zijn van bepaalde uitdagingen om met algoritmen om te gaan, waardoor doorwerking in beleidsvorming soms nog lastig is.

De doelstelling van dit kwalitatieve onderzoek is om kennis te vergroten en theorie te ontwikkelen over de wijze waarop percepties van legitimiteit van ambtenaren de doorwerking beïnvloeden van algoritmen in beleidsvorming. De focus op het concept legitimiteit is relevant omdat het menselijke actoren zijn die bepalen of en hoe algoritmische modellen worden gebruikt. Om dit te onderzoeken is een casestudy uitgevoerd naar het Wmo-voorspelmodel dat door verschillende Nederlandse gemeenten wordt gebruikt. Hierbij stond de volgende onderzoeksvraag centraal: *“Op welke wijze beïnvloedt de perceptie van legitimiteit van ambtenaren de doorwerking van het Wmo-voorspelmodel in de beleidsvorming in Nederlandse gemeenten?”* Dit voorspelmodel dient om het Wmo-gebruik inzichtelijk te krijgen voor de komende vijf jaar. Ook biedt het inzicht in manieren om het gemeentelijk budget effectiever te besteden. Het model voorspelt aan de hand van landelijke CBS-data het aantal gebruikers en kosten voor zowel de stad als per wijk op basis van ongewijzigd beleid.

De voornaamste onderzoeksmethode betrof semigestructureerde interviews met medewerkers uit verschillende gemeenten die allen betrokken zijn (geweest) bij (de ontwikkeling van) het Wmo-voorspelmodel. Uit de interviews is gebleken dat respondenten het model overwegend legitiem beoordelen, zowel inhoudelijk als in het ontwerpproces. Echter, geven respondenten aan dat zij de lokale toepasbaarheid van de uiteindelijke voorspelling missen. Daarom is de conclusie van dit onderzoek dat percepties van legitimiteit een beperkte invloed hebben op de doorwerking van het Wmo-voorspelmodel in de beleidsvorming, waarbij een positieve beoordeling van legitimiteit de doorwerking ervan niet garandeert. Behalve legitimiteit is namelijk ook bruikbaarheid van het model belangrijk om doorwerking ervan te bewerkstelligen. Een van de aanbevelingen is daarom om het doel van het voorspelmodel te herzien met het oog op de behoeftes en doelstellingen van de gemeente, zodat deze beter op elkaar aansluiten.

Inhoudsopgave

1. INLEIDING	5
1.1 AANLEIDING	5
1.2 PROBLEEMSTELLING.....	6
1.3 MAATSCHAPPELIJKE EN BESTUURLIJKE RELEVANTIE	7
1.4 WETENSCHAPPELIJKE RELEVANTIE.....	8
1.5 LEESWIJZER	9
2. ACHTERGROND.....	10
3. THEORETISCH KADER.....	11
3.1 DE DOORWERKING VAN ALGORITMEN IN BELEIDSVORMING	11
3.1.1 <i>Beleidsvorming</i>	11
3.1.2 <i>Informatie ecologie</i>	12
3.1.3 <i>Algorithmic governance</i>	13
3.1.4 <i>Doorwerking</i>	15
3.1.5 <i>Mogelijk beïnvloedende factoren in de doorwerking van algoritmen in beleidsvorming ...</i>	16
3.2 DEFINITIES VAN LEGITIMITEIT IN DE WETENSCHAPPELIJKE LITERATUUR	18
3.2.1 <i>Definitie legitimiteit</i>	18
3.2.2 <i>Deelprocessen van legitimiteit</i>	19
3.3 CONCEPTUEEL MODEL	21
3.4 OPERATIONALISERING.....	23
4. METHODEN.....	25
4.1 ONDERZOEKSSTRATEGIE	25
4.2 ONDERZOEKSMETHODEN	26
4.3 WAARBORGING VAN VALIDITEIT EN BETROUWBAARHEID	28
4.3.1 <i>Betrouwbaarheid</i>	28
4.3.2 <i>Validiteit</i>	28
5. RESULTATEN.....	30
5.1 CASUS	30
5.2 DE DOORWERKING VAN HET WMO-VOORSPELMODEL IN BELEIDSVORMING.....	33
5.2.1 <i>Doorwerking in beleidshandelingen</i>	34
5.2.2 <i>Doorwerking in beleidsdocumenten</i>	35
5.2.3 <i>Sub-conclusie</i>	36
5.3 DE PERCEPTIES VAN LEGITIMITEIT OVER HET WMO-VOORSPELMODEL	37

5.3.1 <i>Input-legitimiteit</i>	37
5.3.2 <i>Throughput-legitimiteit</i>	39
5.3.3 <i>Output-legitimiteit</i>	41
5.3.4 <i>Feedback-legitimiteit</i>	44
5.3.4.3 <i>Sub-conclusie</i>	46
6. CONCLUSIE EN DISCUSSIE	48
6.1 CONCLUSIE	48
6.2 DISCUSSIE	50
6.3 VERVOLGONDERZOEK.....	53
7. REFLECTIE.....	54
8. AANBEVELINGEN.....	56
8.1 AAN GEMEENTEN.....	56
8.2 AAN DE VNG.....	57
BIBLIOGRAFIE	58
BIJLAGE 1 OBSERVATIESCHEMA PARTICIPERENDE OBSERVATIE	63
BIJLAGE 2 RESPONDENTENLIJST	64
BIJLAGE 3 INTERVIEWPROTOCOL.....	65
BIJLAGE 4 DOCUMENTENLIJST	67
BIJLAGE 5 DOCUMENTANALYSE-PROTOCOL	68
BIJLAGE 6 INITIËLE CODEBOOM.....	69
BIJLAGE 7 FINALE CODEBOOM	70

1. Inleiding

1.1 Aanleiding

De doorwerking van algoritmen in het beleidsproces is groeiende, wat het debat over mogelijke implicaties belangrijker dan ooit maakt (Latzer & Festic, 2019). De Algemene Rekenkamer (2021) definieert een algoritme als volgt: ‘een algoritme is een set van regels en instructies die een computer geautomatiseerd volgt bij het maken van berekeningen om een probleem op te lossen of een vraag te beantwoorden’ (p. 11). De uitkomsten die algoritmen geven, kunnen binnen het beleidsproces als informatie en/of kennis dienen ter onderbouwing voor het te vormen beleid (Bowen & Zwi, 2005). Zo kunnen algoritmen bijdragen aan het gericht inzetten van mensen en middelen in bedrijfsvoerings- en dienstverleningsprocessen van organisaties (Algemene Rekenkamer, 2021). Binnen deze vorm van data-gedreven beleid is het belangrijk dat de doorwerking van dit soort technologieën eerlijk en rechtmatig is (Wolfsen, 2021). Het gebruik van algoritmen in beleidsprocessen kan namelijk directe gevolgen hebben voor burgers. Hieraan gerelateerd is de notie van legitimiteit van algoritmen. Legitimiteit gaat over het accepteren van een genomen besluit door een gezaghebbende actor (Airaksinen, Härkönen & Haveri, 2013). Een besluit is legitiem als het zowel legaal als ‘juist’ is (Bekkers, 2017). Wat als ‘juist’ wordt beschouwd is echter lastig te bepalen, omdat dit per context en per actor kan verschillen (Merad & Trump, 2018). In dit onderzoek wordt legitimiteit gekoppeld aan een algoritmisch model, waarmee niet alleen naar het model zelf wordt gekeken, maar ook naar het gebruik ervan.

Algoritmen kunnen in het beleidsproces op verschillende manieren doorwerken. Dankzij het toegenomen vermogen om enorme hoeveelheden data te verzamelen, produceren, verwerken en combineren, zijn geavanceerdere berekeningen en scenario’s mogelijk (Moody & Bekkers, 2022). Deze ontwikkeling draagt bij aan ons begrip en vermogen om te leren van huidige – en toekomstige – beleidsproblemen, omdat de technologie helpt bij het herkennen van betekenisvolle patronen. Zo kunnen we maatschappelijke processen, trends in de samenleving en menselijke interacties beter begrijpen (Moody & Bekkers, 2022). Het meebewegen met deze ontwikkeling, zo wordt gesteld in het whitepaper van DigiTime (2021), vergt zowel een digitale transformatie, als een menselijke. Het whitepaper behandelt de stand van zaken van de Nederlandse overheid op het gebied van digitale transformatie (DigiTime, 2021). Zo lijkt de uitdaging in de praktijk namelijk het overgaan van de drempel van het willen, naar het daadwerkelijk kunnen en doen. Beleidsmedewerkers zijn zich wel bewust van de kansen van nieuwe technologieën, echter weten ze (nog) niet precies wat dat inhoudt of maken

ze zich zorgen om de legitimiteit van het gebruik van data (Moody, 2021). Beleidsmedewerkers bezitten niet altijd de nodige technologische kennis, wat bepaalde barrières creëert (DigiTime, 2021).

Het gebrek aan kennis bemoeilijkt het werken met algoritmen in beleidsvorming, omdat men de uitlegbaarheid van besluiten omtrent beleidsvorming belangrijk vindt (Issar & Aneesh, 2021). Kennis van bijvoorbeeld de manier waarop een algoritmisch model tot een uitkomst is gekomen, is noodzakelijk om te besturen en om bepaalde systemen verantwoordelijk te houden (Selbst & Barocas, 2018). Een veelvoorkomende kritiek op algoritmische besluitvorming is dat het ondoorzichtig en ondoorgrondelijk is. De mate van complexiteit van algoritmen belemmert namelijk de uitlegbaarheid van een algoritmisch model (Janssen & Kuk, 2016). Dit zou het voor mensen lastig maken om de interne werking van het besluitvormingsproces te begrijpen (Selbst & Barocas, 2018).

Men spreekt vanwege de ondoorzichtigheid en ondoorgrondelijkheid vaak van een *black box* (Selbst & Barocas, 2018; WRR, 2021). Echter, wil een algoritmisch model de complexiteit van de realiteit representeren, gaat dat ten koste van de uitlegbaarheid. Zo stellen ook Van den Braak en Choenni (2019): ‘hoe beter een model de wereld representeert, hoe complexer het is om te interpreteren’ (p. 24). In hoeverre een algoritmisch model daadwerkelijk een *black box* is of als zodanig wordt ervaren, maakt het beoordelen van algoritmen in termen van legitimiteit een lastig maar belangrijk proces (Van Veenstra & Kottering, 2017). Die beoordeling bepaalt namelijk of een algoritmisch model legaal en juist wordt geacht, en dus of het geoorloofd is om dit te gebruiken of niet. Zoals eerder benoemd, dient de uitkomst van een algoritmisch model ter onderbouwing om overwogen besluiten te maken. Transparant gebruik van overtuigend bewijs en gepast overleg zijn hierbij essentieel (Head, 2015). Wanneer dit grondig gebeurt, draagt dit bij aan een gebalanceerd en legitiem bestuur. De percepties van betrokken actoren binnen deze processen spelen hierin een belangrijke rol (Moody & Bekkers, 2022). Zij bepalen namelijk welke keuzes en overwegingen over het gebruik van algoritmen worden gemaakt.

1.2 Probleemstelling

Naar aanleiding van bovenstaande is het doel van dit onderzoek om kennis te vergroten en theorie te ontwikkelen over de wijze waarop percepties van legitimiteit van ambtenaren de doorwerking beïnvloeden van algoritmen in de beleidsvorming. Om dit doel te realiseren, is een casestudy uitgevoerd naar de doorwerking van het Wmo-voorspelmodel binnen verschillende Nederlandse gemeenten. Wmo is de afkorting voor de Wet Maatschappelijke

Ondersteuning waarvoor gemeenten verantwoordelijk zijn. De casus zal in Hoofdstuk 5 nader worden toegelicht.

Dit onderzoek is vanwege de schaarste in kennis over dit specifieke onderwerp verkennend van aard (Van Thiel, 2014). Dit betekent dat de resultaten een gedetailleerde empirische omschrijving geven van de casus, waarbij ook aandacht is voor de manier waarop theoretische concepten in de praktijk door actoren worden ervaren. Het onderzoeksdoel zal worden vervuld door de volgende onderzoeksvraag te beantwoorden:

“Op welke wijze beïnvloedt de perceptie van legitimiteit van ambtenaren de doorwerking van het Wmo-voorspelmodel in de beleidsvorming in Nederlandse gemeenten?”

In het theoretisch kader komen de volgende deelvragen aan bod:

1. Wat is bekend in de wetenschap over doorwerking van algoritmen in beleidsvorming?
2. Hoe wordt legitimiteit gedefinieerd in de wetenschappelijke literatuur?

De empirische hoofdstukken behandelen de volgende deelvragen:

1. Op welke wijze is sprake van doorwerking van het Wmo-voorspelmodel in de beleidsvorming in Nederlandse gemeenten?
2. Wat zijn de percepties van legitimiteit van ambtenaren over het Wmo-voorspelmodel in Nederlandse gemeenten?
3. Wat is de relatie tussen de percepties van legitimiteit van ambtenaren over het Wmo-voorspelmodel en de doorwerking ervan in beleidsvorming in Nederlandse gemeenten?

1.3 Maatschappelijke en bestuurlijke relevantie

In de Nederlandse praktijk is momenteel veel aandacht voor (behoorlijk) datagebruik door overheden voor beleidsvorming. Transparantie en privacy zijn daarbij erg belangrijk (Muller, 2021). Recentelijk heeft de Algemene Rekenkamer (2021) een toetsingskader opgesteld om gebruik van algoritmen door de overheid beter te kunnen beoordelen. Uit het whitepaper van DigiTime (2021) blijkt echter dat de menselijke ontwikkeling achterblijft bij de digitale ontwikkeling, terwijl de wens naar (meer) data-gedreven werken wel degelijk aanwezig is. Om mee te kunnen gaan in deze ontwikkeling is het belangrijk om ervaringen uit de praktijk te onderzoeken, hiervan te leren en erop te anticiperen. Dit onderzoek maakt inzichtelijk op welke aspecten winst te behalen valt voor (verdere) doorwerking van het Wmo-voorspelmodel. Dit is ook relevant voor nader te ontwikkelen algoritmische modellen voor het openbaar bestuur. Een

beter vermogen onder ambtenaren om data-gedreven te werken, kan de positie van deze medewerkers versterken (Steunenberg, 2018). Zo kunnen bijvoorbeeld beleidskeuzes in de beleidsvorming beter worden onderbouwd (Van den Braak & Choenni, 2017). Dit is bovendien voor de maatschappij belangrijk, omdat hiermee het beleidsproces responsiever kan zijn naar maatschappelijke patronen, beleidseffecten en behoeften van burgers (Van der Voort, Klievink, Arnaboldi & Meijer, 2019).

Een concrete bijdrage van dit onderzoek is dat de bevindingen van dit verkennende onderzoek als input zullen dienen voor het evaluerende onderzoek naar het landelijke Wmo-voorspelmodel door de Vereniging van Nederlandse Gemeenten (VNG). Op basis van dat onderzoek zal de VNG het model evalueren om besluiten te nemen over structurele voortzetting en eventuele aanpassingen aan het model (VNG, n.d.-1). Zodoende kan de VNG reflecteren op deze bevindingen, op het voorspelmodel zelf en de beoogde doelstellingen om vervolgens het eigen onderzoek richting te geven. Voor gemeenten is dit relevant, omdat hiermee duidelijk wordt of het voorspelmodel helpt bij het beter oplossen van maatschappelijke problemen omtrent de Wmo.

1.4 Wetenschappelijke relevantie

Naast de achterblijvende ontwikkeling van de mens om met de digitale ontwikkeling mee te gaan, wordt in de wetenschap geclaimd dat de beleidscyclus niet snel genoeg wordt doorlopen (Craglia, Hradec & Troussard, 2020). In plaats van dat beleid de snelle ontwikkelingen in de samenleving bijhoudt en leidt, loopt beleid achter de feiten aan. Hier heerst momenteel een kennislacune. Zo stellen wetenschappers zich de vraag hoe we, naast het besturen *van* digitale transformatie, ook ons voordeel kunnen doen door te besturen *met* deze transformatie (Craglia et al., 2020). Vanwege het toenemende gebruik van algoritmen door overheidsinstanties, is het belangrijk om de rol van deze technologie in de samenleving – en in het openbaar bestuur – beter te begrijpen (Levy, Chasalow & Riley, 2021). De toenemende beschikbaarheid van data over beleidseffecten, heeft implicaties voor het beleidsproces, zoals beleidsformuleringen (Steunenberg, 2018).

De doorwerking van algoritmen verschilt echter significant per context. Een empirische benadering naar de doorwerking van een algoritmisch model in de beleidsvorming is daarom relevant, waarbij rekening wordt gehouden met de contextuele en contingente eigenschappen van algoritmen (Kitchin, 2017). Zo wordt bovendien empirisch bewijs aangeleverd voor de gesuggereerde mogelijkheden van algoritmische modellen in beleidsvorming, zoals verbeterde

informatie en verbeterde besluiten (Van der Voort et al., 2019). De focus op het concept legitimiteit is relevant omdat het menselijke actoren zijn die bepalen of en hoe algoritmische modellen worden gebruikt. Percepties van legitimiteit beïnvloeden namelijk het vermogen van een overheid om beoogde doelen te behalen (Wallner, 2008).

1.5 Leeswijzer

Het volgende hoofdstuk zal de relatie uiteenzetten tussen algoritmen en andere in deze context veelgebruikte technologische termen, zoals *machine learning* en artificiële intelligentie. Hoofdstuk 3 behandelt het theoretisch kader, waarna het methodologische hoofdstuk volgt. Hoofdstuk 5 zal de casus van het Wmo-voorspelmodel introduceren om vervolgens de resultaten en analyse van de data te bespreken. Hoofdstuk 6 behandelt de conclusie van het onderzoek door antwoord te geven op de hoofdvraag, waarna tevens een verbinding wordt gemaakt tussen de theorie en empirie. Hoofdstuk 7 reflecteert op sterke en zwakke punten van het onderzoek. Het laatste hoofdstuk bespreekt een aantal aanbevelingen die zijn gebaseerd op de bevindingen.

2. Achtergrond

Voordat de doorwerking van algoritmen kan worden besproken, is een uiteenzetting van het begrip algoritme zelf van belang. Termen als algoritmen, *machine learning* en artificiële intelligentie worden vaak door elkaar gebruikt (Levy et al., 2021). In deze thesis zal, in navolging van Leeuw (2019), de volgende wetenschappelijke definitie van algoritmen worden gebruikt: ‘een lijst van stappen/instructies om met een computerprogramma een probleem op te lossen’ (p. 4). Simpel gezegd zijn algoritmen een vorm van intelligentie die data-input transformeren naar een gewenste uitkomst (Gritsenko & Wood, 2022; Janssen & Kuk, 2016; Kitchin, 2017). Op deze manier vallen zowel simpele als geavanceerde statistische analyses onder de noemer ‘algoritmen’ (Henman, 2020).

De complexiteit neemt toe naarmate algoritmen zelflerend zijn, wat betekent dat de algoritmen leren en voorspellingen maken van grote hoeveelheden data door patroonherkenning (Leeuw, 2020). Dit wordt ook wel *machine learning* (ML) genoemd. Een voorbeeld hiervan is het trainen van algoritmen voor gezichtsherkenning. Ter illustratie: afbeeldingen van gezichten worden als data-input gebruikt voor algoritmen om patronen in te herkennen. Hierop passen algoritmen een regel toe, waarna de output van de algoritmen gezichten kan herkennen uit verschillende afbeeldingen. De volgende stap in complexiteit is artificiële intelligentie (AI). Met AI wordt vaak verwezen naar ML algoritmen die autonoom hun interne regels bepalen om tot een uitkomst te komen (Henman, 2020; Leeuw, 2020). Het verschil tussen ML en AI is dat AI zélf zijn eigen complexe logica ontwikkelt om data te verwerken (Henman, 2020). AI lijkt door zijn vermogen om (tot op zekere hoogte) autonoom te kunnen waarnemen, handelen en zich aan te passen steeds meer te beschikken over menselijke vaardigheden, zoals plannen en leren (Leeuw, 2019).

Over het onderscheid tussen algoritmen, ML en AI kan het volgende worden gesteld: algoritmen zijn niet altijd ML en AI, maar andersom zijn algoritmen wel per definitie onderdeel van ML en AI. De overeenkomst tussen deze technologieën is dat ze gebruikmaken van data en datasets die in verschillende databases en -netwerken worden verwerkt en bewaard (Moody & Bekkers, 2022). Algoritmen kunnen dus variëren in complexiteit, van simpele statistische analyses tot geavanceerde zelflerende ML-technieken die hun eigen interne regels bepalen om output te genereren (Henman, 2020). Naarmate algoritmen complexer worden, wordt tevens het besluitvormingsproces van algoritmen ondoorzichtiger. Dit bemoeilijkt vervolgens de uitlegbaarheid van de uitkomst (Henman, 2020). Zo zijn algoritmen in staat om complexe taken uit te voeren die verder reiken dan het menselijk vermogen (Kitchin, 2017).

3. Theoretisch kader

Dit hoofdstuk behandelt het theoretisch kader door de twee theoretische deelvragen per paragraaf te beantwoorden. Paragraaf 3.1 gaat over wat in de wetenschap bekend is over doorwerking van algoritmen in beleidsvorming. De tweede paragraaf gaat over hoe legitimiteit wordt gedefinieerd in de wetenschappelijke literatuur. Dit hoofdstuk sluit af door in paragraaf 3.3 het conceptueel model te bespreken dat voortvloeit uit de twee voorafgaande paragrafen.

3.1 De doorwerking van algoritmen in beleidsvorming

Allereerst wordt het begrip beleidsvorming uiteengezet, om vervolgens de stap te zetten naar de relatie tussen mens en technologie. Deze stap is nodig om het onderzoek in perspectief te plaatsen, namelijk in die van *algorithmic governance*. De paragraaf sluit af met een verkenning van het begrip doorwerking.

3.1.1 Beleidsvorming

Beleidsvorming staat in het teken van het ontwikkelen van interventies om maatschappelijke problemen aan te pakken (Kolkman, 2020). In de literatuur wordt beleidsvorming veelal benaderd als deelproces van de beleidscyclus (Höchtel, Parycek & Schöllhammer, 2016; Pencheva, Esteve & Mikhaylov, 2020). De beleidscyclus bestaat uit meerdere deelprocessen die door auteurs verschillend worden gedefinieerd. Toch zijn ze vaak te herleiden naar agendavorming, beleidsvorming, -uitvoering en -evaluatie (Bekkers, 2017). Dit onderzoek richt zich op de fase van beleidsvorming. De literatuur stelt dat beleidsvorming bestaat uit beleidsvoorbereiding en beleidsbepaling (ook wel besluitvorming genoemd) (Bekkers, 2017). Deze fase begint met het opstellen van het te ontwerpen beleid in een beleidsprogramma, waarover men vervolgens een besluit neemt om het al dan niet te bekrachtigen (Bekkers, 2017).

Om de fase van beleidsvorming te kunnen realiseren, zijn een aantal tools beschikbaar voor de zogenoemde beleidsanalyse (Kolkman, 2020). Er zijn vragenlijsten en checklijsten, maar ook formele tools zoals kosten-batenanalyse en risico-assessments (Nilsson et al., 2008). Ten slotte zijn er meer geavanceerde tools die de maatschappelijke of economische ontwikkeling in kaart brengen middels op computer gebaseerde modellen en simulaties. Beleidsvorming waarbij data (algoritmen) worden gebruikt, valt onder die laatste categorie. Men spreekt dan ook wel van data-gedreven beleidsvorming (Van Veenstra & Kottering, 2017). Wanneer men beleid onderbouwt met bewijs, afkomstig van bijvoorbeeld experimenteel onderzoek, spreekt men van *evidence-informed* beleid (Bowen & Zwi, 2005). Op basis van de

categorisering van Bowen en Zwi (2005) zou data-gedreven beleid een vorm van *evidence-informed* beleid zijn. Een van de typen bewijs die kunnen worden gebruikt, zijn namelijk kennis en informatie, waaronder ook statistische analyses van algoritmische modellen vallen. Van Veenstra en Kottering (2017) stellen echter dat *evidence-informed* en data-gedreven beleid wezenlijk van elkaar verschillen. Zo ligt bij data-gedreven beleid expliciet de nadruk op het betrekken van vele open databronnen. Dit is anders dan *evidence-informed* beleid, omdat daar ook gebruikgemaakt kan worden van typen bewijs afkomstig van allerlei andere bronnen (Bowen & Zwi, 2005).

Dit onderzoek is gericht op de doorwerking van algoritmen en zal daarom spreken over data-gedreven beleid, in plaats van *evidence-informed*. Zoals in Hoofdstuk 2 is benoemd, maken algoritmen gebruik van data-input om tot een gewenste uitkomst te komen (Gritsenko & Wood, 2022; Janssen & Kuk, 2016; Kitchin, 2017). Data-gedreven beleidsvorming omvat het gebruikmaken van informatie- en communicatietechnologie (ICT) om de voordelen van nieuwe databronnen te kunnen benutten (Van Veenstra & Kottering, 2017). Van Veenstra en Kottering (2017) benoemen daarbij ook dat samenwerking met stakeholders en het betrekken van burgers belangrijke aspecten van data-gedreven beleidsvorming is. Hoewel samenwerking wel zal worden meegenomen, wordt het aspect van het betrekken van burgers achterwege gelaten. Van Veenstra en Kottering (2017) spreken in dat verband namelijk over het opzetten van Policy Labs, waarvan geen sprake is in dit huidige onderzoek.

3.1.2 Informatie ecologie

Een belangrijke factor die het gebruik van (geavanceerde) algoritmen mogelijk maakt, is de beschikbaarheid van enorme hoeveelheden gevarieerde data die snel gegenereerd en verwerkt kunnen worden (Moody & Bekkers, 2022; Pencheva et al., 2020). Dit fenomeen wordt ook wel Big Data genoemd. In de literatuur zijn verschillende perspectieven op Big Data, en technologie in het algemeen, te herkennen (Kitchin, 2017; Moody & Bekkers, 2022). Deze perspectieven duiden de relatie tussen mens en technologie. Die relatie gaat bijvoorbeeld over de autonomie en neutraliteit van de technologie, maar ook over de relatie met de samenleving. Enerzijds wordt technologie gezien als neutraal en waarde vrij, waarbij de mens betekenis en waarde hecht aan de technologie (Kitchin, 2017; Moody & Bekkers, 2022). Anderzijds heeft technologie juist de autonomie om de samenleving te sturen en zitten waarden en moraliteit ingebed in de technologie.

Een ander perspectief, wat de informatie ecologie wordt genoemd, focust niet zozeer op wie of wat de autonomie heeft, maar meer op de context waarin de technologie zich bevindt

(Kitchin, 2017; Moody & Bekkers, 2022). Hier wordt de nadruk gelegd op gemaakte keuzes. De manier waarop een technologie wordt gebruikt en doorwerkt, komt tot stand na bepaalde (politieke en/of sociale) afwegingen van de mens (Moody & Bekkers, 2022). Hierin behoudt de mens dus nog enige autoriteit over de technologie, maar wordt tevens rekening gehouden met een bepaalde mate van beïnvloedingskracht van de technologie. Dit perspectief zal in dit onderzoek worden gehanteerd vanwege twee redenen. Ten eerste, omdat de nadruk niet ligt op wie of wat uiteindelijk de autonomie heeft, maar op het proces dat de doorwerking van de technologie in verschillende contexten kan beschrijven. Ten tweede, omdat keuzes en afwegingen ook van belang worden geacht bij het beoordelen van de legitimiteit van het algoritmische model, het gebruik en de doorwerking ervan.

Vanuit dit opzicht kan technologie enkel binnen de specifieke maatschappelijke en politieke context worden begrepen, waarbij aandacht is voor de interacties en relaties tussen verschillende actoren, praktijken en waarden (Moody & Bekkers, 2022). In lijn met het informatie ecologie perspectief stellen Bowen en Zwi (2005) dat bewustzijn van de context waarbinnen bewijs – zoals data – wordt gebruikt, essentieel is voor effectieve beleidsvorming. Voordat deze contextuele factoren worden besproken, is het van belang om het huidige betoog toe te spitsen op algoritmische modellen in het openbaar bestuur.

3.1.3 *Algorithmic governance*

De alomtegenwoordigheid van algoritmen in het openbaar bestuur maakt dat sommige wetenschappers spreken van *algoracities* in een algoritmisch tijdperk, of van *algorithmic governance* (Gritsenko & Wood, 2022; Kitchin 2017; Latzer & Festic, 2019; Moody & Bekkers, 2022). Van *algorithmic governance* (AG) is sprake wanneer algoritmen samen met andere (traditionele) bestuursystemen de samenleving sturen, bijvoorbeeld door middel van algoritmische selectiesystemen (Latzer & Festic, 2019). Algoritmen fungeren in deze bestuursvorm als tool of als (semi-)autonome agenten. Het gaat hier dus niet om het reguleren/besturen van algoritmen, maar om algoritmen als tool voor reguleren/besturen (Gritsenko & Wood, 2022). In het kader van de informatie ecologie is AG een complex en onderling verbonden systeem van verdeelde *agency*/autonomie tussen mens en technologie (Latzer & Festic, 2019).

Deze bestuursvorm verandert de manier waarop overheden omgaan met maatschappelijke problemen en daarmee ook de manier waarop beleid wordt afgestemd op die op te lossen problemen (Moody & Bekkers, 2022). Voor beleidsvorming specifiek kunnen algoritmen bijvoorbeeld worden ingezet om maatschappelijke trends te scannen, om

beleidsproblemen te voorspellen, om impactanalyses en/of effectmetingen uit te voeren (Leeuw, 2019). De keuze voor het gebruikmaken van algoritmische modellen voor beleidsvorming heeft daarom belangrijke implicaties voor de probleemdefiniëring. Dit geldt voor zowel de aard van het probleem als de uitkomsten van het gevormde beleid (Levy et al., 2021). De manier waarop het algoritmische model wordt ontworpen, reflecteert bovendien veelal de institutionele logica's en representeert de ideeën en overtuigingen van de ontwerpers en beleidsmakers (Janssen & Kuk, 2016). Bij het ontwerpen van een algoritmisch model wordt namelijk toegewerkt naar een gewenst resultaat, waarbij ook ongewenste effecten – zoals etnisch profileren – worden geprobeerd tegen te gaan.

Zodra het algoritmische model is ontworpen en ingezet, ontstaat de vraag welke rol de mens nog dient te spelen, bijvoorbeeld in het beoordelen van het algoritmische besluit. Deze tussenkomst van de mens in het besluitvormingsproces van algoritmen wordt ook omschreven als *humans in the loop* (De-Artega, Fogliato & Chouldechova, 2020). Hoewel er discussie bestaat over de wenselijkheid van *humans in the loop*, worden algoritmen in de praktijk gebruikt om alledaagse besluiten te ondersteunen ofwel te verbeteren door middel van aanbevelingen, filteren of het aanbrengen van scores (De-Artega et al., 2020; Latzer & Festic, 2019; Levy et al., 2021). Het besluitvormingsproces wordt zelden tot nooit volledig vervangen door de technologie. In de praktijk hebben algoritmen dus niet de volledige autoriteit om besluiten te nemen, maar behouden beleidsmedewerkers de discretie om algoritmische besluiten te verwerpen (Levy et al., 2021). De mate waarin de mens de discretie behoudt, is van belang om in ogenschouw te houden, omdat dit behulpzaam is in het evalueren van AG in termen van betrouwbaarheid en verantwoording (Latzer & Festic, 2019).

De alomtegenwoordigheid van algoritmische modellen maakt de doorwerking ervan daarentegen niet vanzelfsprekend. De visie van Kitchin (2017) op algoritmen is hierin verhelderend: '*algorithms need to be understood as relational, contingent, contextual in nature, framed within the wider context of their socio-technical assemblage*' (p. 18). Hiermee wordt bedoeld dat algoritmen ontstaan door processen als discussie en onderhandeling, wat maakt dat ze altijd in ontwikkeling zijn door continue aanpassingen en herzieningen (Janssen & Kuk, 2016; Kitchin, 2017). Het sociaal-technische aspect verwijst enerzijds naar de technische kenmerken (de software), anderzijds naar de sociale kenmerken (het gebruik door mensen) die algoritmen omvatten (Latzer & Festic, 2019). Wederom is hier het informatie ecologie perspectief te herkennen. Zo stelt Kitchin (2017) dat met een focus op de lokale context waarin algoritmen zich begeven, kennis, organisatie- en institutionele cultuur, wettigheden en persoonlijke overtuigingen belangrijke factoren zijn om rekening mee te houden. Een

belangrijke vraag, als het gaat om de doorwerking van algoritmen in beleidsvorming, richt zich bijvoorbeeld op welke afwegingen worden gemaakt, gebaseerd op persoonlijke overtuigingen door betrokken actoren, evenals heersende ideeën en voorkeuren binnen de organisatie (Moody & Bekkers, 2022).

3.1.4 Doorwerking

In de literatuur wordt met het concept doorwerking vaak de doorwerking van beleid bedoeld (De Lange, 1995). Dit is in het huidige onderzoek niet het geval. In plaats daarvan wordt gedoeld op de doorwerking van algoritmische modellen in de beleidsvorming. Zo kan doorwerking, in navolging van De Lange (1995), als volgt worden gedefinieerd: ‘de nadere uitwerking, meer concrete invulling, vertaling en toepassing van [algoritmen] door beleidsactoren die hierop al dan niet formeel worden aangesproken’ (p. 15).

Om dit begrip analytisch te kunnen onderzoeken, wordt in de wetenschap onderscheid gemaakt tussen twee benaderingen (De Lange, 1995). Doorwerking kan enerzijds als proces worden benaderd, anderzijds als referentiekader. Met referentiekader wordt bedoeld dat de nadere uitwerking van initieel beleid binnen bepaalde grenzen moet vallen. Het initiële beleid is in die zin richtinggevend, waarbij doorwerking volgens De Lange (1995) meer als beoordelingscriterium wordt gebruikt voor de effectiviteit van beleid. Hoewel beide benaderingen analytisch interessant zijn, wordt in dit onderzoek enkel het proces van doorwerking geanalyseerd. Het meten van het effect van doorwerking vraagt namelijk om bewijs waaruit blijkt dat een bepaald gevolg toe te wijzen is aan het gebruik van bepaalde informatie (Rich, 1997). Dat is echter niet de vraag die hier wordt gesteld. Waar het wel om gaat is de rol die percepties van legitimiteit spelen in het al dan niet gebruikmaken van algoritmische modellen. Gezien deze focus is de benadering van doorwerking als proces meer gepast. Zoals De Lange (1995) en Rich (1997) stellen, zijn daarbij de volgende vragen van belang: waar is het doorwerkingsproces op gericht, wie zijn daarbij betrokken en wat is het doel? Dit vraagt om een gedetailleerde beschrijving van het proces, waaruit tevens verschillen in doorwerking kunnen blijken (De Lange, 1995).

Door doorwerking als proces te benaderen, wordt aandacht gegeven aan de verschillende rollen die de technologie feitelijk speelt, welke het zou moeten spelen of voor welke rol het in eerste instantie bedoeld was (De Lange, 1995). Potman (in De Lange, 1995, p. 38) onderscheidt bovendien verschillende typen doorwerking. Ten eerste kan doorwerking plaatsvinden in beleidsdocumenten van aangesproken actoren. Met actoren worden in dit onderzoek de gebruikers van algoritmische modellen bedoeld die het model raadplegen. De

beleidsdocumenten waar De Lange (1995) naar refereert zijn binnen de context van doorwerking veelal strategisch van aard. Dit houdt in dat beleid voor de langere termijn is bedoeld. Strategische beleidsdocumenten krijgen daarom vaak de vorm van een beleidsplan. Ten tweede kan sprake zijn van doorwerking in operationele, dagelijkse beslissingen en handelingen van deze actoren.

Omdat verschillende gebruikers betrokken zijn bij het doorwerkingsproces, is het inzichtelijk om hierin een onderscheid in te maken. Er zijn initiators van de technologie, maar ook overige actoren die in het verloop van het doorwerkingsproces worden aangesproken. Deze overige actoren worden ook wel nageschakelde actoren genoemd (De Lange, 1995). Gezien deze verschillende posities in het doorwerkingsproces, kunnen van initiators andere percepties, keuzes en overwegingen worden verwacht in vergelijking met de nageschakelde actoren. Zodoende is een verdeling te maken naar typen doorwerking en typen actor. Deze verdeling is te zien in Tabel 1.

Tabel 1

Doorwerking gecategoriseerd naar type doorwerking en type actor

Doorwerking in beleidsdocumenten Vanuit initiator	Doorwerking in beleidsdocumenten Vanuit nageschakelde actor
Doorwerking in operationele, dagelijkse beslissingen en handelingen Vanuit initiator	Doorwerking in operationele, dagelijkse beslissingen en handelingen Vanuit nageschakelde actor

Opmerking: Aangepast overgenomen uit De Lange (1995, p. 39).

3.1.5 Mogelijk beïnvloedende factoren in de doorwerking van algoritmen in beleidsvorming

De idee dat data onderdeel zouden moeten zijn van de *toolkit* van bestuurders is groeiende (Verhulst, Engin & Crowcroft, 2019). Toch blijkt in de praktijk dat het gebruik van Big Data methoden en analyses nog niet van de grond komt (Poel, Meyer & Schroeder, 2018). Regelmatig blijft het bij geïsoleerde prototypen, trials of pilots (Verhulst et al., 2019). Zo wordt binnen de overheid momenteel nog relatief weinig gebruikgemaakt van zeer complexe algoritmen als AI, in tegenstelling tot de minder geavanceerde algoritmen zoals verfijnde statistische analyses (Henman, 2020).

De contextafhankelijkheid van het gebruik van algoritmen vergroot het belang van zorgvuldig toegespitste beleidsbesluiten (Latzer & Festic, 2019). Zo brengt een algoritmische selectie voor het aanraden van een boek andere risico's met zich mee dan eenzelfde algoritmisch

model dat een medische behandeling aanraadt. Gebruikers van algoritmische modellen dienen met zulke verschillende risico's rekening te houden. Een van de risico's hangt samen met de complexiteit van het algoritmische model (Van den Braak & Choenni, 2017). Naarmate het model in complexiteit toeneemt, zullen mensen minder in staat zijn om het functioneren te begrijpen (Janssen & Kuk, 2016). Met name ML algoritmen en AI worden gekenmerkt door hun ondoorgrondelijkheid en ondoorzichtigheid (Selbst & Barocas, 2018). Voor deze geavanceerde algoritmen geldt dat het *black box*-gehalte het begrijpelijk, veilig, transparant en geloofwaardig houden van het model bemoeilijkt (Leeuw, 2020). Een besluit dat met behulp van AI is ondersteund, maakt bovendien het verantwoordingsproces lastiger, omdat een algoritmisch besluit niet de menselijke *agency*/autonomie heeft (Henman, 2020; Janssen & Kuk, 2016). Dit maakt de vraag over legitimiteit en de perceptie daarop van belang, wat in de volgende paragraaf aan bod komt.

Dat de doorwerking van algoritmen vragen oproept omtrent uitlegbaarheid, verantwoording en transparantie, betekent volgens sommige wetenschappers echter niet dat de technologie overboord moet worden gegooid (Craglia et al., 2020). AG biedt namelijk ook mogelijkheden die het beleidsvormingsproces kunnen ondersteunen. Een van de mogelijkheden is dat het voor beleidsmakers steeds makkelijker wordt om data-gedreven beleid te vormen (Van den Braak & Choenni, 2017). Zo kunnen data ter onderbouwing van beleid of ter ondersteuning van de onderhandelingspositie dienen (Kolkman, 2020). Een andere kans wordt personalisatie genoemd (Levy et al., 2021). Personalisatie verwijst naar het gebruik van geaggregeerde data waarop beleidsplannen worden gebaseerd. Dit maakt het categoriseren van burgers, toewijzen van diensten en voorspellen van gedrag op individueel niveau mogelijk. Vermeende voordelen hiervan zijn dat de toewijzing van middelen verbetert, het kosten vermindert en preventieve interventie aanwakkert (Levy et al., 2021). Een ander wijdverspreid idee van optimisme is dat algoritmen de kwaliteit van besluiten verhoogt door preciezere voorspellingen, waarmee prestaties verbeteren (De-Artega et al., 2020). Desalniettemin benadrukken wetenschappers dat het van belang blijft om menselijke intuïtie en beredenering in het besluitvormingsproces te behouden, omdat de kans altijd bestaat dat een voorspelling het mis heeft (Selbst & Barocas, 2018). Hier wordt gewezen op de notie van *humans in the loop*, zoals besproken in sub-paragraaf 3.1.3 over AG.

Zoals eerder benoemd, is dit onderzoek niet expliciet geïnteresseerd in de verhouding tussen mens en technologie. Vanuit de informatie ecologie zijn de gemaakte keuzes en overwegingen in de lokale context ook interessant (Moody & Bekkers, 2022). Dit wordt in dit

onderzoek benaderd vanuit de notie van legitimiteit, wat het onderwerp is van de volgende paragraaf.

3.2 Definities van legitimiteit in de wetenschappelijke literatuur

Deze paragraaf is opgedeeld in twee secties. De eerste sub-paragraaf bespreekt de definitie van legitimiteit waarna dit begrip wordt gekoppeld aan algoritmische modellen. De tweede sub-paragraaf benoemt verschillende deelprocessen waar legitimiteit tot uiting kan komen. Ook deze concepten wordt vervolgens gekoppeld aan algoritmische modellen.

3.2.1 Definitie legitimiteit

In de literatuur wordt legitimiteit op verschillende manieren benaderd. Veelal richt de bestuurskundige literatuur zich op de legitimiteit van het publieke beleid (Wallner, 2008). In plaats van beleid, draait het in dit onderzoek om de legitimiteit van het algoritmische model. Hierin wordt niet enkel gekeken naar het ontwerp van het model, maar ook naar het resultaat dat het model geeft, de omgang met en doorwerking van dat resultaat, en de herziening van het model. Wanneer het gaat over de percepties van legitimiteit, wordt gedoeld op de percepties onder betrokken ambtenaren bij het algoritmische model in de beleidsvorming. Omdat de algemene bestuurskundige notie van legitimiteit inzichtelijk is, wordt deze behandeld alvorens deze toe te passen op algoritmische modellen.

Bij het begrip legitimiteit staat het beoordelen van de uitoefening van macht centraal, ofwel het nemen van (bindende) besluiten (Bekkers, 2017). Zoals Hoofdstuk 1 aanstipte, is een besluit legitiem wanneer het zowel legaal als ‘juist’ is (Bekkers, 2017). Het wetmatige wijst op het uitoefenen van gezag dat formeel ligt vastgelegd in wet- en of regelgeving (Van Buuren, Driessen, Teisman & Van Rijswick, 2014). Belangrijk hierin is dat overheden reflecteren op hun besluiten en deze kunnen onderbouwen. De juistheid van een besluit benadrukt meer het normatieve aspect van legitimiteit (Scharpf, 2009). In zijn geheel kan legitimiteit vanuit twee perspectieven worden beoordeeld, namelijk inhoudelijk en procedureel (Jagers, Matti & Nordblom, 2016; Van Buuren et al., 2014). Ook Wallner (2008) volgt deze tweedeling en zal tevens in dit onderzoek worden aangehouden. Hierbij zullen de twee aspecten zoals hierboven benoemd, namelijk de wetmatigheid en juistheid, wel in ogenschouw worden gehouden, maar niet expliciet worden opgenomen in de operationalisering.

De inhoudelijke beoordeling van legitimiteit kijkt naar de mate waarin wordt gehandeld naar de verschillende normen, waarden, belangen en overtuigingen van de gemeenschap

(Bekkers, 2017; Van Buuren et al., 2014). Vanuit een procedurele bril gaat het om de kwaliteit van het doorlopen besluitvormingsproces (Bekkers, 2017). Dit verwijst bijvoorbeeld naar de manier waarop gezag wordt uitgeoefend naar uniform geldende en logisch consistente regels zoals in een (democratische) rechtsstaat mag worden verwacht (Bekkers, 2017). Belangrijke aspecten hierbij zijn verantwoording, de stem van de burger en zorgvuldig overleg (Van Buuren et al., 2014). Onder zorgvuldig overleg wordt verstaan dat sprake is van transparantie en toegang tot informatie. De notie van transparantie houdt in dat gezaghebbende besluiten open toegankelijk zijn voor de betrokken actoren (De Fine Licht, 2011). Voor alle aspecten waarop legitimiteit kan worden beoordeeld, geldt echter dat dit afhangt van iemand zijn waardeoordeel van bijvoorbeeld wat juist is, welke democratische waarden van belang zijn en wat andere essentiële voorwaarden zijn (Van Buuren, Klijn & Edelenbos, 2012).

Toegepast op algoritmen houdt dit in dat een algoritmisch model legitiem is – in het ontwerp, de gegeven uitkomst, het gebruik en de herziening – wanneer het past binnen de normen en wettelijke kaders vanuit de perceptie van de betrokken actoren. In dit onderzoek wordt dit geanalyseerd vanuit de percepties van ambtenaren, omdat het de percepties zijn waarop actoren hun handelen baseren, zingeven en evalueren (De Lange, 1995). Omdat inhoudelijke legitimiteit niet automatisch procedurele legitimiteit verzekert, is het wenselijk om vanuit beide perspectieven tot een positieve beoordeling van legitimiteit te komen (Bekkers, 2017). De algehele beoordeling in termen van legitimiteit is belangrijk, omdat dit het vermogen van een overheid beïnvloedt om beoogde doelen te halen (Wallner, 2008).

3.2.2 Deelprocessen van legitimiteit

Naast de twee verschillende niveaus om legitimiteit te beoordelen (procedureel en inhoudelijk), is ook een verdeling in deelprocessen van legitimiteit behulpzaam, omdat tevens voor het begrip doorwerking de procesbenadering wordt gevolgd. De literatuur onderscheidt de volgende deelprocessen van beleidsvorming: input-, *throughput*- output-, en feedback-legitimiteit (Bekkers, 2017). Binnen deze deelprocessen kan legitimiteit op verschillende manieren tot uiting komen. Wederom gaat deze indeling niet om de beleidsvorming zoals veelal in de literatuur is bedoeld, maar past dit onderzoek legitimiteit toe op een algoritmisch model. Zoals in de vorige sub-paragraaf is gesteld, wordt niet enkel gekeken naar het ontwerp, maar ook naar de gegeven voorspelling, het omgaan met die voorspelling, de doorwerking en de herziening van het model. Deze opdeling van het algoritmische model wordt in dit onderzoek gekoppeld aan de vier deelprocessen van legitimiteit.

Input-legitimiteit gaat over de mate waarin de voorkeuren, wensen en belangen van de betreffende gemeenschap in acht zijn genomen (Airaksinen et al., 2013; Bekkers, 2017). De focus ligt op de posities die voorafgaand aan het besluitvormingsproces worden ingenomen (Van Buuren et al., 2012). Gekoppeld aan het algoritmische model worden hier de data bedoeld die als input dienen voor het ontwerp van het model, alsmede de ingenomen posities tijdens het ontwikkelingsproces. Legitimiteit in het verloop van het proces wijst op de *throughput*-legitimiteit, zoals het volgen van bepaalde procedures (Bekkers, 2017). Zo zullen verschillende procedures verschillende normen en waarden beschermen. Zorgvuldig overleg tijdens collaboratieve processen om bijvoorbeeld beleidsproblemen te analyseren of te definiëren zijn voorbeelden die de *throughput*-legitimiteit kunnen bevorderen (Van Buuren et al., 2012). Dit onderzoek verwijst met *throughput*-legitimiteit naar het proces (het algoritme) waarmee het model tot een resultaat komt.

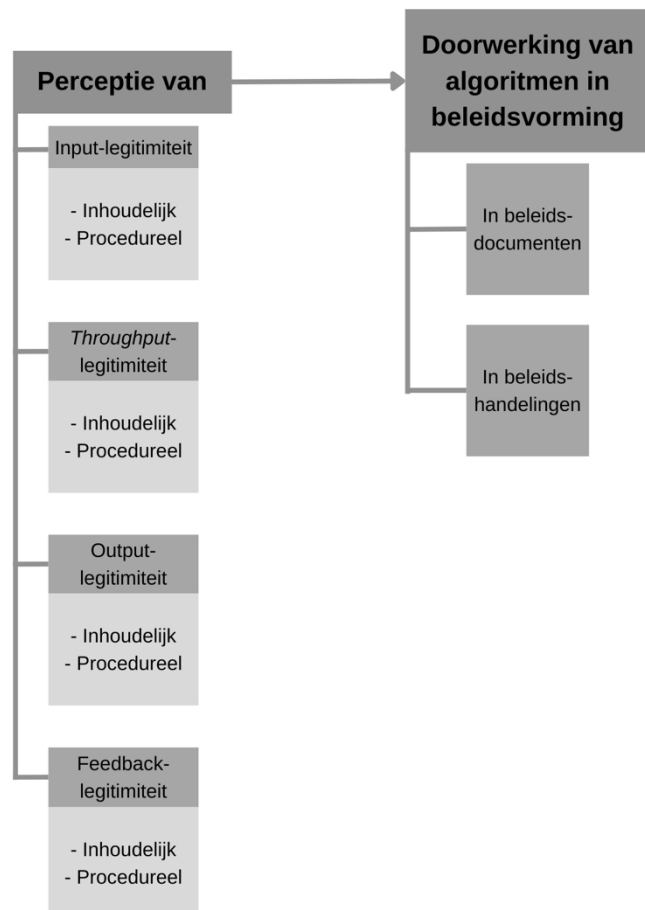
Output-legitimiteit focust op het al dan niet accepteren van een resultaat (Airaksinen et al., 2013). Om output-legitimiteit te bevorderen dienen bijvoorbeeld de gevolgen van een besluit gunstig te zijn voor de betrokkenen (Jagers et al., 2016). Omdat dit onderzoek gaat over een voorspelmodel, wordt hier gesproken van een voorspelling in plaats van resultaat of besluit. Dat is dan ook waar output-legitimiteit naar verwijst; de voorspelling zoals gegeven door het model. Ten slotte, feedback-legitimiteit. Dit gaat over de vraag in hoeverre het systeem omgaat met de resultaten van een besluit of beleidsinterventie door ze terug te koppelen als nieuwe input of ter reflectie (Bekkers, 2017). Feedback-legitimiteit wordt verbonden aan het terugkoppelen van verbeterpunten, adviezen, op- en aanmerkingen wat betreft het model zelf en/of het gebruik ervan. In navolging van Jagers et al. (2016) suggereert het onderscheid in verschillende deelprocessen dat de perceptie van legitimiteit per proces kan variëren. Bovendien zijn percepties cultuur- en contextafhankelijk, wat maakt dat bestaande politieke en sociale processen in ogenschouw genomen dienen te worden, wat wederom past binnen het informatie ecologie perspectief (Airaksinen et al., 2013).

3.3 Conceptueel model

Op grond van bovenstaande uiteenzettingen van de doorwerking van algoritmen in de beleidsvorming en de perceptie van legitimiteit, kan een conceptueel model worden opgesteld zoals te zien is in Figuur 1. Aan de hand van dit model, zullen ook de antwoorden op de twee theoretische deelvragen beknopt worden behandeld. Omdat het model van links naar rechts wordt gelezen, zal eerst paragraaf 3.2 over legitimiteit worden samengevat.

Dit onderzoek focust op de menselijke actoren die bepalen of en hoe een algoritmisch model doorwerkt. Daarom zijn het de percepties van betrokken ambtenaren die worden onderzocht. Dit wordt gedaan in termen van legitimiteit, omdat dit begrip het vermogen van een overheid beïnvloedt om beoogde doelen en doelstellingen te halen (Wallner, 2008). Legitimiteit heeft hier betrekking op het algoritmische model, in tegenstelling tot het beleid zoals in veel andere wetenschappelijke literatuur.

Samenvattend uit Paragraaf 3.2 komt legitimiteit, toegepast op algoritmen, neer op het volgende: een algoritmisch model is legitiem – in het



Figuur 1 Conceptueel model

ontwerp, de gegeven uitkomst, het gebruik en herziening – wanneer het past binnen de normen en wettelijke kaders vanuit de perceptie van de betrokken actoren. Legitimiteit kan worden beoordeeld in zowel procedurele als inhoudelijke termen (Wallner, 2008). Inhoudelijk gaat om het handelen naar verschillende normen, waarden, belangen en overtuigingen in de gemeenschap (Bekkers, 2017; Van Buuren et al., 2014). Procedureel wijst op de kwaliteit van het doorlopen besluitvormingsproces (Bekkers, 2017). Omdat voor het begrip doorwerking als proces wordt benaderd, wordt tevens het concept legitimiteit onderverdeeld in verschillende deelprocessen, namelijk de input-, output, *throughput*-, en feedback-processen (Jagers et al., 2016).

Uit het bovenstaande model is op te maken dat er een vermeende relatie bestaat tussen percepties van legitimiteit en de doorwerking van algoritmen in beleidsvorming. Samenvattend uit paragraaf 3.1 bestaat beleidsvorming uit enerzijds beleidsvoorbereiding, waarbij men beleid ontwerpt in een beleidsprogramma, en beleidsbepaling anderzijds, wat inhoudt dat men het beleid al dan niet bekrachtigt (Bekkers, 2017). Vanuit het perspectief van de informatie ecologie is beleidsvorming waarin algoritmen een rol spelen een complex en onderling verbonden systeem, waarbij de autonomie tussen mens en technologie is verdeeld (Moody & Bekkers, 2022). In de praktijk betekent dit dat beleidsvorming waarbij algoritmen worden gebruikt, de mens nog altijd algoritmische besluiten kan verwerpen (Levy et al., 2021). Dit is belangrijk, omdat naarmate een algoritmisch model complexer wordt, de uitlegbaarheid van het algoritmische besluit vermindert (Janssen & Kuk, 2016). De complexiteit van beleidsvorming maakt tevens dat het onderscheid tussen beleidsvoorbereiding en beleidsbepaling in de praktijk vaak niet dusdanig strikt is, maar dat deze eerder chaotisch of cyclisch van aard is (Bekkers, 2017).

Om toch de manier waarop algoritmische modellen worden gebruikt in beleidsvorming te kunnen doorgronden, wordt het begrip doorwerking gebruikt. Doorwerking is de nadere uitwerking, meer concrete invulling, vertaling en toepassing van algoritmen door beleidsactoren die hierop al dan niet formeel worden aangesproken (De Lange, 1995). Met de procesbenadering van doorwerking is onderscheid gemaakt tussen zowel de typen doorwerking als de typen actoren (Potman in De Lange, 1995, p. 38). Enerzijds onderscheidt de literatuur het type doorwerking in beleidsdocumenten, anderzijds het type in operationele, dagelijkse beslissingen en handelingen. De typen actoren zijn te onderscheiden in de initiator van de technologie, en de zogeheten nageschakelde actoren.

Hoewel de idee dat algoritmen van significante toegevoegde waarde kunnen zijn in het openbaar bestuur, is de doorwerking in de praktijk niet vanzelfsprekend (Poel et al., 2018). De keerzijde van een complex maar kwalitatief hoogwaardig algoritmisch model dat de samenleving juist representeert, gaat namelijk ten koste van de begrijpelijkheid en transparantie van het model (Leeuw, 2020). Toch zijn sommige wetenschappers optimistisch over de doorwerking van algoritmen in beleidsvorming, omdat het bijvoorbeeld de kwaliteit van besluiten kan verhogen doordat preciezere voorspellingen van maatschappelijke trends mogelijk zijn (De-Artega et al., 2020). Het probleem van voorspellen is echter dat de kans bestaat dat de voorspelling ernaast zit, wat de discussie omtrent *humans in the loop* zo van belang maakt (Selbst & Barocas, 2018). Het is de perceptie van ambtenaren op deze en andere facetten in termen van legitimiteit waar dit onderzoek in is geïnteresseerd.

3.4 Operationalisering

Om alle concepten uit het model te kunnen meten, zijn de concepten geoperationaliseerd (Van Thiel, 2014). De indicatoren geven aan hoe de concepten kunnen worden herkend in de empirie. Tabel 2 operationaliseert de afhankelijke variabele ‘doorwerking van algoritmen in beleidsvorming’ en Tabel 3 de onafhankelijke variabele ‘de perceptie van legitimiteit’. Per deelproces van legitimiteit is zowel de inhoudelijke als procedurele dimensie aangegeven.

Tabel 2

Afhankelijke variabele 'doorwerking van algoritmen in beleidsvorming'

Concept	Typen	Indicatoren
Algoritmen	Simpele statistische analyses	Een algoritmisch model maakt statistische berekeningen middels bijvoorbeeld beslisbomen
	<i>Machine learning</i>	Een algoritmisch model leert en maakt voorspellingen van grote hoeveelheden data door patroonherkenning
	Artificiële intelligentie	Een algoritmisch model bepaalt autonoom zijn eigen interne regels om tot een uitkomst te komen
<hr/>		
Concept	Typen	Indicatoren
Doorwerking van algoritmen in beleidsvorming	Doorwerking in beleidsdocumenten	De voorspelling van het algoritmisch model wordt gebruikt in beleidsdocumenten
	Doorwerking in beleidshandelingen	De voorspelling wordt geraadpleegd voor operationele, dagelijkse beslissingen en handelingen

Tabel 3

Onafhankelijke variabele 'de perceptie van legitimiteit'

Concept	Dimensies	Indicatoren
Input- legitimiteit: data die als input voor het model worden gebruikt	Inhoudelijk	De gebruikte data-input representeren de juiste normen, waarden, belangen en overtuigingen
	Procedureel	Het doorlopen proces om tot de keuze voor bepaalde data-input te komen is van voldoende kwaliteit
<i>Throughput</i> : de onderliggende algoritmen waarmee het model tot een voorspelling komt	Inhoudelijk	De onderliggende algoritmen waarmee het model tot een voorspelling komt, zijn juist
	Procedureel	Het doorlopen proces om tot de keuze voor bepaalde algoritmen te komen, is van voldoende kwaliteit
Output: de voorspelling die het model geeft	Inhoudelijk	De gegeven voorspelling is juist
	Procedureel	Het omgaan met de voorspelling is van voldoende kwaliteit
Feedback: terugkoppeling van verbeterpunten, adviezen, op- en aanmerkingen	Inhoudelijk	De genoemde verbeterpunten, adviezen, op- en aanmerkingen zijn inhoudelijk juist
	Procedureel	Het omgaan met de gegeven feedback-punten is van voldoende kwaliteit

4. Methoden

Dit hoofdstuk behandelt de methodologische verantwoording en uitvoering van het onderzoek. Allereerst zal de onderzoeksstrategie worden uiteengezet. De tweede paragraaf stelt welke maatregelen zijn genomen om de validiteit en betrouwbaarheid van het onderzoek te waarborgen.

4.1 Onderzoeksstrategie

Om de hoofdvraag te beantwoorden is een kwalitatief onderzoek uitgevoerd. Een kwalitatief onderzoek werd als meest gepast beoordeeld, omdat de interesse lag in de percepties van ambtenaren. De sociale realiteit zoals die door ambtenaren wordt ervaren en omschreven, was hetgeen waar dit onderzoek naar op zoek was, wat een kwalitatieve benadering logisch maakte (Boeije, 2010). Uit vooronderzoek bleek dat in de praktijk de wens heerst om data-gedreven te werken, maar dat sprake is van verschillende obstakels om die wens te realiseren. Zo zijn er bijvoorbeeld zorgen over de legitimiteit van het gebruik van data of is de technologische kennis beperkt (DigiTime, 2021; Moody, 2021). Dit riep de vraag op welke rol percepties van legitimiteit hier mogelijk spelen.

Voorafgaand aan het empirische onderzoek is een uitgebreide literatuurstudie gedaan, om inzichtelijk te krijgen wat al dan niet wetenschappelijk bekend is over het onderwerp algoritmen in beleidsvorming en legitimiteit (Babbie, 2016; Van Thiel, 2014). Aangezien geen theorie voor handen lag die een relatie aantoonde tussen de perceptie van legitimiteit en de doorwerking van algoritmen in beleidsvorming, is dit onderzoek hoofdzakelijk inductief van aard (Van Thiel, 2014). Dit houdt in dat op basis van specifieke observaties in de empirie algemene principes worden geformuleerd (Babbie, 2016). Zo hebben de empirische observaties tot een theoretisch mechanisme geleid. Omdat het vormen van een gehele theorie meer onderzoek vergt, heeft dit onderzoek zich beperkt tot het opstellen van een theoretisch mechanisme dat als bouwsteen kan dienen voor de nader te vormen theorie. Hoewel het onderzoek hoofdzakelijk inductief, heeft het ook deductieve kenmerken, omdat de concepten uit het theoretisch kader het vervolg van het onderzoek hebben aangestuurd.

Voor een casestudy-benadering is gekozen, omdat deze vorm geschikt is om data te verzamelen en te analyseren van variabelen die complex en niet-numeriek zijn, zoals percepties (Van Thiel, 2014). Een casestudy focust op een enkel geval van een bepaald sociaal fenomeen (Babbie, 2016). Het Wmo-voorspelmodel werd als geschikte casestudy geacht, omdat het een voorbeeld is van een algoritmisch model dat in beleidsvorming kan doorwerken. Om het

onderzoek ingekaderd te houden, zijn gemeente-medewerkers geïnterviewd die betrokken zijn bij een eigen en/of het landelijke Wmo-voorspelmodel. Met een doelbewuste steekproef zijn respondenten geselecteerd die kennis van de casus hadden (Van Thiel, 2014). Met de meeste respondenten is via Janneke Lummen, projectleider namens VNG Realisatie, contact gelegd, omdat zij als centraal contactpersoon het beste kon beoordelen wie geschikte respondenten waren. De respondenten waren afkomstig uit vijf verschillende gemeenten, namelijk: Roosendaal, Rotterdam, Den Haag, Deventer en Velsen. Deze set aan gemeenten was bevorderlijk voor de verscheidenheid aan perspectieven op het Wmo-voorspelmodel, omdat zij in verschillende mate betrokken waren bij de ontwikkeling van het landelijke model. Zo waren de gemeenten Roosendaal, Rotterdam en Velsen koplopers in de ontwikkeling en was gemeente Den Haag adviserend betrokken (Punte, 2021). Gemeente Deventer sloot in een latere fase aan bij het ontwikkeltraject. De diversiteit aan perspectieven komt niet enkel door verschil in betrokkenheid, maar ook door verschil in omvang.

4.2 Onderzoeksmethoden

Er zijn meerdere onderzoeksmethoden gebruikt. Ter voorbereiding op het daadwerkelijke onderzoek zijn in februari t/m april 2022 kennismakende gesprekken gevoerd met academici en medewerkers bij verschillende overheden die kennis hebben van algoritmen binnen het openbaar bestuur. Zo is bekend geraakt met de context van het onderzoek (Babbie, 2016). Met hetzelfde doel is participerende observatie (PO) uitgevoerd tijdens de leergang “Omgaan met algoritmen en AI binnen de overheid” van ScienceWorks op 7, 14, 21 en 28 april 2022. PO is een onderzoeksmethode waarbij de onderzoeker onderdeel wordt van het onderzochte fenomeen (Boeije, 2010). De verzamelde data zijn later gebruikt als input voor interviewvragen en als achtergrondkennis over de materie uit de praktijk. Het gebruikte observatieschema is te vinden in Bijlage 1.

De voornaamste methode betrof zeven semigestructureerde interviews waarbij in totaal tien respondenten zijn gesproken. In kwalitatieve onderzoeken hebben interviews doorgaans een relatief open karakter (Babbie, 2016). Bijlage 3 bevat de vooraf opgestelde vragenlijst. Deze methode paste het beste bij het doel om de perceptie van beleidsmedewerkers in hun eigen woorden uitgelicht te krijgen. Er werd namelijk veel waarde gehecht aan de sociale ervaringen, waarden en interpretaties vanuit de respondenten, waardoor het belangrijk was dat zij open, uitgebreid en in hun eigen woorden interviewvragen konden beantwoorden (Babbie, 2016). Het

semigestructureerde karakter maakte ook mogelijk om te reageren op wat respondenten zeiden en om door te vragen.

Om de data te analyseren zijn de interviews (na toestemming van de respondent) opgenomen en getranscribeerd. De interviews vonden plaats tussen eind april en begin juni op Zoom en Teams. Bijlage 2 bevat een (geanonimiseerd) overzicht van de respondenten. Sommige respondenten gaven aan het interview het beste samen met een andere collega te houden, zodat zij vanuit de inhoud- en data-kant de interviewvragen konden beantwoorden. Daarom zijn drie interviews met twee respondenten tegelijk afgenomen. Om te verzekeren dat de interpretatie van de data juist was, is een membercheck gedaan (Van Thiel, 2014). Zo is aan respondenten die daar behoeften aan hadden, het concept-resultatenhoofdstuk opgestuurd, en kregen zij de mogelijkheid te wijzen op foute interpretaties. Ook zijn vier (beleids)documenten geanalyseerd om eventuele expliciete doorwerking van het Wmo-voorspelmodel waar respondenten over spraken inzichtelijk te krijgen. Zo waren de onderzoeksmethoden complementair aan elkaar. Vanwege gemaakte afspraken om als gemeente niet herleidbaar te zijn, kan niet worden gespecificeerd welke documenten zijn meegenomen in de analyse. Naast gemeentelijke documenten is tevens een document van de VNG geanalyseerd, omdat dit de ontwikkeling van het landelijke Wmo-voorspelmodel verhelderde. Bijlage 4 toont een overzicht van de geanalyseerde documenten, Bijlage 5 het gevolgde protocol.

De verzamelde data zijn vervolgens gestructureerd, geanalyseerd en geïntegreerd door ze allereerst gesloten te coderen (Van Thiel, 2014). Dit wijst op de deductieve eigenschap van dit onderzoek, omdat de gebruikte codes corresponderen met de operationalisering. De initiële gesloten codeboom staat in Bijlage 6. Citaten die iets zeiden over de theoretische concepten werden onder de betreffende code gevoegd. Ook is gebruikgemaakt van memo's om waar nodig extra toelichting te geven bij een codering en/of interpretatie van data (Van Thiel, 2014). In de tweede fase van het coderen werd de gesloten codeboom uitgebreid met thematische en beschrijvende codes om de citaten te groeperen. Gedurende dit proces van axiaal coderen zijn stapsgewijs patronen ontdekt die mogelijke correlaties aantonen tussen de variabelen (Van Thiel, 2014). Dit proces had een iteratief karakter doordat de data werden geanalyseerd terwijl de dataverzameling voortduurde. Zo is heen en weer gegaan tussen het verzamelen en analyseren van de data. De finale codeboom staat in Bijlage 7.

4.3 Waarborging van validiteit en betrouwbaarheid

4.3.1 Betrouwbaarheid

Volgens Van Thiel (2014) is een onderzoek betrouwbaar als de metingen van de variabelen zowel accuraat als consistent zijn. Dit verzekert dat de resultaten op systematische wijze zijn verkregen en niet willekeurig zijn. De mate van accuraatheid van een onderzoek wijst op de gebruikte meetinstrumenten die precieze en correcte metingen mogelijk maken (Van Thiel, 2014). Om de accuraatheid van dit onderzoek te borgen, zijn alle interviews opgenomen en getranscribeerd. Dit bevordert de precisie van de data-analyse. Ook de keuze voor het afnemen van interviews is voor dit aspect bevorderlijk. Respondenten konden in hun eigen woorden hun percepties omschrijven, wat een preciezere en accurate omschrijving van de bevindingen mogelijk maakte.

Consistentie houdt in dat het onderzoek herhaalbaar is, ofwel dat bij herhaling dezelfde resultaten uitkomen (Neuman, 2014; Van Thiel, 2014). Dit onderzoek is echter vanwege de kwalitatieve benadering interactief en situationeel van aard, wat maakt dat het verkrijgen van identieke resultaten bij herhaling niet gegarandeerd is (Neuman, 2014). Hierdoor is de betrouwbaarheid van het onderzoek enigszins beperkt. Om dit te compenseren zijn de gebruikte methoden transparant en compleet uiteengezet. Zo is het onderzoek navolgbaar gemaakt. Het coderen in Atlas.ti heeft gezorgd voor een overzicht van de uitgevoerde analyse. Daarbij is gebruikgemaakt van memo's die een nadere toelichting geven bij de interpretatie van de data. Dit ondersteunt niet alleen de interpretatie van de data, maar ook de navolgbaarheid ervan.

Ten slotte is de algehele betrouwbaarheid geborgd dankzij de betrokkenheid van een scriptiebegeleider gedurende het gehele proces. De feedbackmomenten zijn daarbij gunstig geweest voor het juist voortzetten van het onderzoek, waarbij fundamentele fouten zijn voorkomen (Van Thiel, 2014). Hoofdstuk 7 reflecteert of bovenstaande maatregelen het doel om de betrouwbaarheid te borgen hebben gerealiseerd.

4.3.2 Validiteit

Ook de validiteit is op meerdere manieren gepoogd te waarborgen. Validiteit kan worden onderscheiden in interne en externe validiteit (Van Thiel, 2014). Interne validiteit gaat erover dat gemeten wordt wat gemeten diende te worden. Dit is geborgd doordat ten eerste een uitgebreid literatuuronderzoek is uitgevoerd om abstracte termen te conceptualiseren en te operationaliseren (Neuman, 2014). Een grondige verdieping in de literatuur maakte heldere formuleringen mogelijk van wat wel en niet werd verstaan onder de te onderzoeken concepten.

Deze maatregel diende bovendien om de content validiteit te verhogen door de volledige definitie van de concepten te dekken (Neuman, 2014). Met het operationaliseren zijn indicatoren geïdentificeerd die valide metingen van de concepten en onderlinge relaties in de empirie mogelijk maakten (Neuman, 2014). Bovendien werd tijdens het coderen het operationaliseringsschema continu ernaast gehouden, om te verzekeren dat de fragmenten correct werden gecodeerd.

De PO en kennismakende gesprekken dienden ter ondersteuning bij het vertalen van abstracte ideeën naar empirisch observeerbare indicatoren. De bedoeling was om bekend te raken met de gebruikte woordenschat in het veld, maar ook met verschillende concrete toepassingen van algoritmen in de beleidscyclus. Door bekend te raken met de materie in de praktijk, werd het wederzijdse begrip met respondenten verbeterd. Zo werd gepoogd om zo dicht mogelijk bij de authentieke betekenissen en uitleg van participanten te blijven (Neuman, 2014). Een andere maatregel was de selectie van een casestudy. De keuze voor het Wmo-voorspelmodel concretiseerde de in- en uitsluitingscriteria van participanten voor het onderzoek. Deze selectiecriteria verzekerden dat enkel de perceptie van gemeentemedewerkers die betrokken zijn (geweest) bij het Wmo-voorspelmodel werden meegenomen. De specificiteit van de casestudy kan enerzijds de externe validiteit belemmeren, omdat de resultaten enkel voor deze casus gelden en dus moeilijk te generaliseren zijn. Anderzijds kan dit juist de interne validiteit versterken (Van Thiel, 2014). Bovendien is externe validiteit niet het doel van dit onderzoek, maar is een holistische benadering nagestreefd door de focus te leggen op een enkele casus.

5. Resultaten

Dit hoofdstuk bespreekt de empirische bevindingen uit dit onderzoek. Allereerst wordt de casus beschreven, om vervolgens de resultaten te presenteren over de doorwerking van het Wmo-voorspelmodel in de beleidsvorming. Daarna komen resultaten over de percepties van legitimiteit van ambtenaren over het model aan bod. Beide resultaten-paragrafen zullen eindigen met een sub-conclusie.

5.1 Casus

Gemeenten zijn sinds de decentralisatie in 2015 verantwoordelijk voor de ondersteuning van mensen die niet op eigen kracht zelfredzaam zijn (Movisie, 2013; Rijksoverheid, n.d.). Dit staat vastgelegd in de Wet maatschappelijke ondersteuning, die officieel Wmo 2015 heet (Rijksoverheid, n.d.). De toegang tot ondersteuning, zoals begeleiding en dagbesteding, regelt iedere gemeente op zijn eigen manier. Middels de Wmo zorgen gemeenten ervoor dat mensen zo lang mogelijk thuis kunnen blijven wonen (Rijksoverheid, n.d.). Burgers doen steeds meer beroep op de Wet maatschappelijke ondersteuning (De Jong, 2019). In 2018 begon de gemeente Den Haag met het ontwikkelen van een voorspelmodel om het Wmo-gebruik inzichtelijk te krijgen voor de komende vijf jaar, en om het gemeentelijke budget effectiever te besteden (Gemeente Den Haag, 2020; iBestuur, 2022).

Het model voorspelt aan de hand van landelijke (open) CBS-data het aantal gebruikers en kosten voor zowel de stad als per wijk op basis van ongewijzigd beleid. Het model dient om grip te krijgen op toekomstige ontwikkelingen, waaronder de uitgaven in het sociaal domein. Het Wmo-gebruik wordt in totaal voorspeld in unieke Wmo-gebruikers, maar ook op voorzieningenniveau. De voorzieningen zijn ingedeeld in drie categorieën: Hulp bij het Huishouden, Ondersteuning Thuis en Hulpmiddelen & Diensten (Gemeente Den Haag, 2021). Het model voorspelt voor de komende vijf jaar en is daarmee gericht op de langere termijn om beleid en uitvoering te ondersteunen op tactisch en strategisch niveau. Het voorspelmodel is een regressiemodel waarbij gebruik wordt gemaakt van kenmerken die het beste het Wmo-gebruik kunnen voorspellen (Gemeente Den Haag, 2021). De gebruikte algoritmen maken gebruik van *machine learning* (D3).

In 2020 leverde gemeente Den Haag ook een landelijk model op dat werd overgedragen aan de VNG om door te ontwikkelen voor gebruik door alle Nederlandse gemeenten (VNG Realisatie, 2021). Zoals blijkt uit een intern Miro-bord (D1) van de VNG is deze doorontwikkeling gefaseerd aangepakt, waarbij in fase 1 een aantal koplopergemeenten hebben

samengewerkt. Er werd een Kernteam Inhoud aangesteld dat inhoudelijk vanuit de Wmo het voorspelmodel beoordeelde, door bijvoorbeeld de belangrijkste onderwerpen vast te stellen en het model te beoordelen op toepasbaarheid en landelijke bruikbaarheid. Daarnaast werd een Kernteam Data aangesteld (D1). Dit team verwerkte de input en feedback afkomstig van de inhoud om tot een eerste versie van het voorspelmodel te komen. In de tweede fase werd het model aan een tweede groep gemeenten beschikbaar gesteld die ermee aan de slag konden. Deze groep gaf maandelijks feedback aan de koplopergemeenten en het projectteam van de VNG (D1). Dit leidde tot een landelijk op te schalen model dat op 24 januari 2022 officieel is gelanceerd. Om een beeld te geven van hoe het Wmo-voorspelmodel eruitziet en werkt, worden hieronder een aantal afbeeldingen laten zien. Afbeelding 1 toont de startpagina van het Wmo-voorspelmodel.



Afbeelding 1 Startpagina van het landelijke Wmo-voorspelmodel: Aangepast overgenomen uit *Wmo voorspelmodel* van VNG, n.d. (<https://wmo-voorspelmodel.vng.nl/>). Copyright, n.d., VNG.

Op de startpagina kan de gebruiker per gemeente het aantal inwoners, niveau van stedelijkheid en Socio-economische Status (SES) zien, evenals het aandeel 65-plussers en Wmo-gebruikers. Hoe dit eruit ziet, laat Afbeelding 2 van zowel een grote als kleine gemeente zien.



Afbeelding 2 Informatie van respectievelijk een grote en kleine gemeente: Aangepast overgenomen uit *Wmo voorspelmodel* van VNG, n.d. (<https://wmo voorspelmodel.vng.nl/>). Copyright, n.d., VNG.

Ter illustratie is in Afbeelding 3 gemeente Moerdijk genomen om de grafieken te tonen die de voorspellingen voor de betreffende gemeente aangeven.



Afbeelding 3 Voorspellingen Wmo-gebruikers van gemeente Moerdijk. Aangepast overgenomen uit *Wmo voorspelmodel* van VNG, n.d. (<https://wmo voorspelmodel.vng.nl/>). Copyright, n.d., VNG.

Ten slotte laat Afbeelding 4 een wijkoverzicht zien van gemeente Moerdijk. In dit geval zijn de wijken overeenkomstig met de plaatsnamen binnen de betreffende gemeente.

Wijkoverzicht + Voeg wijkgroep toe - 2020 + Lijst Kaart

Selecteer wijken in de lijst om ze toe te voegen aan de grafiek hierboven

Wijk	Wmo gebruikers	Aandeel binnen de wijk	Inwoners	SES	65-plussers
Wijk 00 Zevenbergen	1.110	7,7%	14.360	gemiddeld	24,0%
Wijk 03 Klundert	450	7,7%	5.865	gemiddeld	20,2%
Wijk 07 Fijnaart	430	7,6%	5.640	laag	22,8%
Wijk 06 Standaardbuiten	110	4,9%	2.240	hoog	18,5%
Wijk 09 Willemstad	110	4,4%	2.510	heel hoog	23,1%
Wijk 01 Zevenbergschen Hoek	90	5,5%	1.635	gemiddeld	19,6%
Wijk 04 Moerdijk	60	5,0%	1.195	gemiddeld	18,4%
Wijk 05 Noordhoek	50	4,6%	1.095	heel hoog	16,4%
Wijk 02 Langeweg	45	5,2%	865	heel hoog	15,6%
Wijk 08 Heijningen	35	4,1%	855	hoog	18,1%
Wijk 10 Helwijk	35	4,0%	880	heel hoog	18,8%

Afbeelding 4 Wijkoverzicht van gemeente Moerdijk. Aangepast overgenomen uit *Wmo voorspelmodel* van VNG, n.d. (<https://wmovoorspelmodel.vng.nl/>). Copyright, n.d., VNG.

Indien de lezer zelf nader het landelijke Wmo-voorspelmodel wil bekijken, kan diegene de website bezoeken zoals aangegeven in de bronvermelding onder de afbeeldingen. Nu is de vraag of het model wordt gebruikt door Nederlandse gemeenten, en indien van toepassing: hoe? Deze vraag is voor dit onderzoek omgezet in een eerste deelvraag, namelijk: op welke wijze is sprake van doorwerking van het Wmo-voorspelmodel in de beleidsvorming in verschillende Nederlandse gemeenten? De bevindingen omtrent deze vraag zijn het onderwerp van de volgende paragraaf.

5.2 De doorwerking van het Wmo-voorspelmodel in beleidsvorming

Deze paragraaf is opgedeeld in drie sub-paragrafen. Allereerst wordt de doorwerking in beleidshandelingen besproken. Sub-paragraaf 5.2.2 bespreekt de doorwerking in beleidsdocumenten. De paragraaf sluit af met een beknopte sub-conclusie.

5.2.1 Doorwerking in beleidshandelingen

Zoals besproken in Hoofdstuk 3 is sprake van doorwerking in beleidshandelingen als het model onderdeel is van operationele, dagelijkse beslissingen en handelingen van actoren (Potman in De Lange, 1995, p. 38). Uit de interviews blijkt dat het Wmo-voorspelmodel niet een instrument is dat onderdeel is van de dagelijkse praktijk, maar wordt geraadpleegd als een specifieke aanleiding ervoor is. Meerdere respondenten geven aan dat ze het model meer zien als iets dat op vaste momenten in het jaar kan worden gebruikt, zoals het opstellen van de programmabegroting of bij coalitieonderhandelingen (R1, 3, 4, 5, 7). Het model dient hier om te kunnen onderbouwen hoe op de te verwachte ontwikkelingen te anticiperen.

Ook richting de politiek kan het model bruikbaar zijn om financiële claims kracht bij te zetten. Bij het opstellen van budgetten is het model bijvoorbeeld behulpzaam wanneer het een bepaalde ontwikkeling aantoont in de zorgvraag (R3 – senior beleidsmedewerker). Het voorspelmodel kwantificeert die ontwikkeling waarop de begroting gebaseerd zou kunnen worden. Een andere aanleiding om het model te raadplegen, is om de voorspelling van het model naast de eigen cijfers te leggen (R1 – beleidsadviseur). Op die manier kunnen de eigen gemaakte prognoses worden vergeleken met die van het voorspelmodel. Wanneer daar grote verschillen in zitten, kan dat aanzetten tot nader onderzoek naar de ontwikkelingen in de Wmo. Zo dient het model meer als tool ter ondersteuning van andere taakstellingen van de respondenten en is het niet een doel op zich (R1, 3). Een concreet voorbeeld wordt gegeven door R3:

“Op het moment dat we een aantal scootmobielplaatsen willen uitbreiden. Ja, dat is het moment dat ik tegen zo’n projectleider zeg van joh, vraag niet alleen aan de individuele wijkteams: wat zou een goede plek zijn? Maar zorg ook dat je het voorspelmodel bekijkt, met alle mitsen en maren.”

Een kanttekening is dat het Wmo-voorspelmodel dat binnen de gemeente van R3 wordt gebruikt, een model is dat enkel ontwikkeld is voor de betreffende gemeente en dus niet het landelijke model is dat de VNG lanceerde. Dit geldt ook voor R2, 4, 7 en 8. Daarbij is het model van de gemeenten van R7 en 8 met een ander doel ontwikkeld dan de andere voorspelmodellen. Het model van R7 en 8 geeft een voorspelling voor een periode van vier weken, in plaats van voor de komende vijf jaar. Een model met een ander doel, vraagt ook om ander gebruik (R8 – data-scientist). R7 (strategisch adviseur) zegt daarover:

“Het liefst zou ik hem elke maand gebruiken, om daar dan mee in gesprek te kunnen gaan met onze contractmanagers van joh, dit zien wij. [...] Kunnen we die zorg nog steeds garanderen met elkaar? [...] Als je het niet zo vaak doet, dan raakt de routine er een beetje uit en dan moet je weer opnieuw gaan kijken, hoe zat het ook alweer?”

Een aanleiding om het model te raadplegen kan ook afkomstig zijn vanuit de uitvoering, zoals de teammanagers (R2 – business- en data-analist). Hoewel dit nog beperkt voorkomt, heeft R2 eens een capaciteitsberekening gedaan naar het aantal klanten en de groei in begeleiding, hulp bij het huishouden en woonvoorzieningen. Zo kan ook de uitvoering rekening houden met de eigen capaciteit en mensen efficiënter inzetten (R2). De reden waarom het Wmo-voorspelmodel niet onderdeel is van de dagelijkse praktijk, is simpelweg omdat het model bedoeld is voor de langere termijn (R3, 4). Zo legt R4 (onderzoeker) uit:

“Je wilt eigenlijk niet weten wat er morgen gebeurt met dit model, maar echt gewoon wat komt er over een jaar of over een jaar of vijf op ons af? En ja, waar kunnen wij dan op anticiperen als gemeente? [...] Dat zijn allemaal langere termijn toepassingen waar dat model voor gebouwd is.”

5.2.2 Doorwerking in beleidsdocumenten

Wat betreft de doorwerking in beleidsdocumenten blijkt uit de interviews dat dit varieert. Zo is in de gemeenten van R2-4, 7 en 8 wel sprake van doorwerking in beleidsdocumenten, maar in gemeenten van R1, 5, 6, 9 en 10 niet. Opvallend is dat daar waar deze doorwerking wél plaatsvindt, gebruik wordt gemaakt van een eigen ontwikkeld Wmo-voorspelmodel, in plaats van het landelijke model. De bruikbaarheid van het model ligt volgens meerdere respondenten met name aan de aan- of afwezigheid van adequate capaciteit en expertise binnen gemeenten (R1, 5, 6). Kleinere gemeenten hebben vaak niet die expertise in huis met voldoende specialistische kennis (R1, 5). Zo is het allereerst van belang dat er kennis is over de technische kant, voordat het model inhoudelijk is te interpreteren voor doorwerking in de beleidspraktijk (R5 – data-analist, 6 – data-engineer). Voor gemeenten waar niet de nodige capaciteit en/of kennis aanwezig is, was deelname aan de ontwikkeling van het landelijke Wmo-voorspelmodel met name gericht op ervaring opdoen (R1, 5, 6). Het aanwakkeren van data-gedreven werken was ook een van de onderliggende ideeën van dit project om gemeenten te laten zien wat de mogelijkheden ervan zijn (R1 – beleidsadviseur). Zo is de betrokkenheid bij het model voor

respondenten an sich al een leerzame ervaring wat bijdraagt aan het vermogen om data-gedreven te werken. R6 zegt hierover:

“Ik heb wel veel ervan geleerd in hoe dit soort dingen werken, en ook zo veel ermee geprutst en tegenaan gelopen dat we daar best wel veel van geleerd hebben. Maar om het nou direct echt toe te kunnen passen in beleidsstukken is gewoon wat lastiger.”

Ondanks dat niet alle gemeenten de expertise en/of capaciteit bezitten voor verdere doorwerking van het model in beleidsvorming, heerst wel degelijk de ambitie om dit door te zetten (R1, 9, 10). Terwijl de een nog de mogelijkheden bekijkt om het voorspelmodel te gebruiken in Wmo-beleidsplannen, heeft de ander concrete voorbeelden van bestuurlijke stukken waarin het model is gebruikt. Voorbeelden van bestuurlijke stukken waarin het Wmo-voorspelmodel is gebruikt, zijn een strategische verkenning, een commissiebrief, en een rapportage waarin een momentopname van de stad wordt gegeven en trends worden besproken (R3 – senior beleidsmedewerker). Het model diende hier veelal om de verwachte toename van het Wmo-gebruik aan te geven. Omdat men zonder voorspelmodel eigenlijk een ‘wilde gok’ doet over verwachte ontwikkelingen omtrent de Wmo, zit de toegevoegde waarde van het model met name in het onderbouwen van die ontwikkelingen (R3, 5). Hierover zegt R3:

“Het is een voorspelling en je hebt bepaalde onzekerheidsmarges et cetera, maar het geeft in ieder geval een onderbouwde indruk van goh, wat verwachten we eigenlijk?”

5.2.3 Sub-conclusie

Op basis van bovenstaande bevindingen kan worden gesteld dat van doorwerking in beleidshandelingen weinig tot geen sprake is. Wat betreft de doorwerking in beleidsdocumenten luidt de conclusie dat dit sterk varieert per context, maar dat de ambitie om het door te ontwikkelen wel aanwezig is. Opvallend is dat daar waar gemeenten gebruikmaken van een eigen model wél sprake is van doorwerking in beleidsdocumenten. In gemeenten die enkel het landelijke voorspelmodel tot hun beschikking hebben, is geen sprake van deze doorwerking. Daarbij wordt een beperkte doorwerking niet meteen afgeschreven als iets negatiefs, maar wordt het belang van de leerzame ervaring om betrokken te zijn bij het Wmo-voorspelmodel als groot voordeel geacht.

5.3 De percepties van legitimiteit over het Wmo-voorspelmodel

Deze paragraaf is opgedeeld zoals ook de operationalisering is verdeeld, namelijk in de deelprocessen van legitimiteit (input, *throughput*, output en feedback). Per deelproces wordt vervolgens de inhoudelijke dimensie beschreven en daarna de procedurele. Ook deze paragraaf sluit af met een sub-conclusie.

5.3.1 Input-legitimiteit

Input-legitimiteit verwijst naar de data die als input voor het model worden gebruikt. Dit is in de inhoud legitiem als de gebruikte data de juiste normen, waarden, belangen en overtuigingen representeren (Airaksinen et al., 2013; Bekkers, 2017; Van Buuren et al., 2014). De input is procedureel legitiem als het doorlopen proces om tot de keuze voor bepaalde data-input van voldoende kwaliteit is (Bekkers, 2017).

5.3.1.1 Inhoudelijke input-legitimiteit

Wanneer de respondenten het hebben over de gebruikte data-input, gebruiken zij veelal de term ‘kenmerken’ of ‘factoren’ (R1-4). Hiermee worden kenmerken bedoeld die een bepaalde mate van voorspelkracht hebben, zoals het aantal 65-plussers. Uit de interviews komen drie aspecten naar voren die van belang zijn voor de inhoudelijke juistheid van de gebruikte kenmerken, namelijk de voorspelkracht, de data-kwaliteit en het doel van het voorspelmodel. Ten eerste, de voorspelkracht. De juistheid van de gebruikte kenmerken voor het landelijke model wordt geborgd door zowel vanuit de data-kant als vanuit de inhoud te beoordelen of de kenmerken een voorspellende waarde hebben (R1, 2). Een voorbeeld is dat de groei van het aandeel 65-plussers samenhangt met het Wmo-gebruik, wat maakt dat je een inschatting kunt maken van de hoeveelheid Wmo-gebruik (R4 – onderzoeker). Zo is het de statistiek die volgens R3 bepaalt welke kernmerken een voorspellende waarde hebben:

“Wie bepaalt dan op basis waarvan je die variabelen kiest? Ja, de statistiek in zekere zin, want een variabele heeft voorspelkracht of niet. Daar zit verder weinig tussen en er zit ook niet een politieke keuze in.”

Het gebruik van de juiste kenmerken is essentieel, omdat dit tevens de kwaliteit van de voorspelling bepaalt, ofwel de inhoudelijke output-legitimiteit (R1, 3, 4). Hieraan verbonden is het tweede aspect, namelijk de data-kwaliteit. Een voorwaarde voor een goed voorspelmodel is

dat de data van voldoende kwaliteit is (R1 – beleidsadviseur). Hoewel dit een open deur lijkt, is die kwaliteit niet voor iedere gemeente vanzelfsprekend, vertelt R1:

“Zo’n voorspelmodel dat valt en staat bij goede kwalitatieve data. Dus het is ook erg afhankelijk van wat gemeenten aanleveren. Ze zagen bijvoorbeeld dat voor sommige gemeenten geen voorspelling gemaakt kon worden, omdat [de data] al een paar jaar of niet waren aangeleverd, of dat ze maar een deel hadden aangeleverd.”

Het derde aspect dat van belang blijkt, is het doel van het voorspelmodel. Zo zitten er significante verschillen tussen het voorspelmodel dat in de gemeente van R7 en R8 wordt gebruikt en in die van de rest. Het voorspelmodel van R7 en R8 is niet bedoeld om te kunnen vergelijken met andere gemeenten, terwijl dat bij de ontwikkeling van het landelijke model wel als een belangrijke functie werd gezien. R7 (strategisch adviseur) ziet daarvan niet de toegevoegde waarde:

“Met Wmo, denk ik, is het vrij lastig om een vergelijking te maken met een andere gemeente. Vind maar eens een gemeente die, bijvoorbeeld, evenveel inwoners heeft [...] en daarbij ook nog een College heeft dat hetzelfde werkt [...] Want soms ga je als College wel andere zwaartepunten neerleggen die dan weer gevolgen hebben voor je beleid.”

Het ontwerpen van een landelijk inzetbaar voorspelmodel heeft belangrijke implicaties gehad voor de lokale bruikbaarheid van het model (R1). Meerdere respondenten geven aan behoefte te hebben aan het verder kunnen inzoomen op voorzieningenniveau (R1, 5, 6, 9, 10). Wat dit lastig maakt is dat niet alle gemeenten dezelfde termen voor voorzieningen gebruiken (R1). De doorwerking van het landelijke voorspelmodel vereist een bepaalde vertaalslag naar de lokale context die door sommige respondenten als lastig wordt ervaren (R9, 10). Tegelijkertijd is er begrip over dit onontkoombare dilemma, omdat een (te) gedetailleerd model zowel de landelijke als lokale bruikbaarheid tenietdoet. Zo zegt R3 (senior beleidsmedewerker):

“Dan ga je natuurlijk onwijs een kerstboom bouwen van zo’n model, en dat is volgens mij heel complex. [...] Volgens mij werkt dat de betrouwbaarheid en de overzichtelijkheid van je model ook niet in de hand.”

5.3.1.2 Procedurele input-legitimiteit

Zoals hiervoor al werd gesteld, is de juistheid van de gebruikte kenmerken in het landelijke model geborgd door zowel het perspectief van data-specialisten te horen als van domein-experts (R1, 2). Het is dit proces van gestructureerde samenwerking die voor respondenten het vertrouwen geeft dat het proces juist is doorlopen om tot de gekozen kenmerken te komen (R1-4). Dit zorgt volgens R1 ervoor dat de kans op mogelijke fouten minimaliseert, zoals het vergeten van een belangrijk voorspellend kenmerk. Het juist doorlopen van dit proces verzekert tegelijkertijd de inhoudelijke juistheid van de kenmerken, doordat wordt vertrouwd op de kennis en expertise van de andere betrokken collega's (R3). R9 (business-analist) spreekt over een transparant doorlopen proces, wat bijdraagt aan het vertrouwen in de juistheid van het voorspelmodel.

Eerder werd al benoemd dat respondenten behoefte hebben aan het verder kunnen inzoomen op voorzieningenniveau. Respondenten uiten meerdere wensen wat betreft de gebruikte data, maar tonen desalniettemin begrip voor het beoordelingsvermogen van collega's, zo blijkt uit de volgende reactie van R3 (senior beleidsmedewerker):

“Je hebt het liefst altijd zo recent mogelijke basisgegevens. Alleen dan is het aan [data-specialist] om op de rem te trappen en te zeggen van ja, maar je kan wel recentere gegevens willen hebben, maar die voorspellen niet goed, of die zijn niet betrouwbaar.”

5.3.2 Throughput-legitimiteit

Throughput-legitimiteit wijst op de onderliggende algoritmen waarmee het model tot een voorspelling komt. Dit is inhoudelijk legitiem als de onderliggende algoritmen als juist worden beoordeeld. Het is procedureel legitiem als het doorlopen proces om tot de keuze voor de onderliggende algoritmen te komen van voldoende kwaliteit is (Bekkers, 2017).

5.3.2.1 Inhoudelijke *throughput*-legitimiteit

Net zoals dat in de keuze voor de juiste data-input wordt vertrouwd op de kennis en expertise van collega's, is dat ook het geval voor de keuze voor de juiste onderliggende algoritmen (R1-3, 5, 9). Die onderliggende algoritmen worden ook wel een 'code' genoemd (R1, 6, 8). Respondenten die zelf niet betrokken waren bij het programmeren van het voorspelmodel vertrouwen erop dat het model goed in elkaar zit, omdat wordt verondersteld dat collega's met data-expertise goed erover hebben nagedacht (R3, 5). Deze respondenten bezitten zelf beperkte

kennis over de werking van de code, maar vinden overigens ook niet dat dit van belang is voor het gebruik van het model (R1-3, 5). Zo legt R1 (beleidsadviseur) uit:

“Ik denk dat het een beetje te ver gaat om te denken dat je het algoritme ook helemaal moet begrijpen voordat ik die voorspelling snap. Ik denk dat het meer is dat je begrijpt wat de methodiek erachter is.”

Wederom komt hier de beperking naar voren voor gemeenten die niet voldoende capaciteit en expertise in huis hebben. Voordat de voorspelling kan worden gebruikt en geïnterpreteerd, moet iemand in staat zijn om het technische gedeelte te begrijpen (R6 – data-engineer). Bij de ontwikkeling van het landelijke model, werd het gebrek aan de nodige ervaring dan ook als beperkend ervaren, blijkt uit wat R6 vertelt:

“Feit blijft dat je hier gewoon best wel wat specialistische kennis voor nodig hebt. Dus het is niet zomaar plug and play. En je merkte ook dat in die bijeenkomsten met name mensen die echt ervaring hadden met Data Science en programmeren in R, dat die duidelijk veel meer kennis ervan hadden en ook vrijer waren om te zeggen ja, maar dit kan beter.”

Een belangrijk kenmerk van zowel het landelijke voorspelmodel als de voorspelmodellen die andere gemeenten zelf hebben ontwikkeld, is dat het geen *black box*-modellen zijn waarbij onduidelijk is wat onder de motorkap gebeurt (R4, 8). Het uitgangspunt om zo transparant mogelijk te zijn over data-gedreven werken was voor de gemeente van R4 bepalend bij de keuze voor een zogeheten *white box*-model. *White box* houdt in dat men weet hoe een voorspelling tot stand is gekomen, welke data zijn gebruikt, en hoe de relatie is tussen de voorspelde aantallen Wmo-gebruikers en de data die nodig zijn om deze voorspelling te kunnen doen (D3).

5.3.2.2 Procedurele *throughput*-legitimititeit

Wanneer erop wordt vertrouwd dat de data-experts de juiste keuzes maken in de onderliggende algoritmen, wordt hiermee ook vertrouwen gegeven in dat het proces juist is doorlopen. Hierbij is goede onderlinge communicatie tussen domein- en data-experts een belangrijk aspect. Meerdere respondenten geven aan vertrouwen te hebben in de zorgvuldigheid van dit ontwerpproces (R2, 3, 5, 9). Zoals eerder is aangegeven, wordt het niet noodzakelijk geacht voor domein-experts om de onderliggende techniek volledig te begrijpen (R1-3, 5). Dit maakt

echter het vertrouwen in, en de expertise van de data-experts extra van belang, zoals R4 (onderzoeker) stelt:

“Er moet voor een deel ook vertrouwen zijn, ook in de mensen die dit soort modellen bouwen, dat ze dat goed hebben gedaan en dat zij alle regels die je moet hanteren, dat ze die hebben gehanteerd en dat dat goed is gelopen.”

Goede onderlinge communicatie met onder andere domein-experts is belangrijk om te beoordelen wat een acceptabele verwachte voorspelfout is (R4). Die communicatie is nodig omdat dit niet alleen contextafhankelijk is, maar ook omdat niet iedereen weet wat bijvoorbeeld vanuit de statistiek acceptabele onzekerheid is (R4, 6). De voorspelfout is het verschil tussen hetgeen dat is voorspeld en hetgeen dat het daadwerkelijk is geworden (R8). Inherent aan voorspelmodellen is dat er altijd een onzekerheidsmarge bestaat (R1-10, D3). Idealiter is de voorspelfout zo klein mogelijk, wat echter een complexer (*black box*) model vereist. Dit is een belangrijke overweging die de gemeente intern bediscussieert (R4 – onderzoeker). Zodoende is – in de ogen van respondenten – een zorgvuldig proces doorlopen om tot een voldoende transparant model te komen met een acceptabele verwachte voorspelfout.

5.3.3 Output-legitimiteit

Output-legitimiteit gaat om de voorspelling die het model geeft. Dit is inhoudelijk legitiem als de voorspelling juist is. Procedureel is dit legitiem als het omgaan met de voorspelling van voldoende kwaliteit is (Bekkers, 2017).

5.3.3.1 Inhoudelijke output-legitimiteit

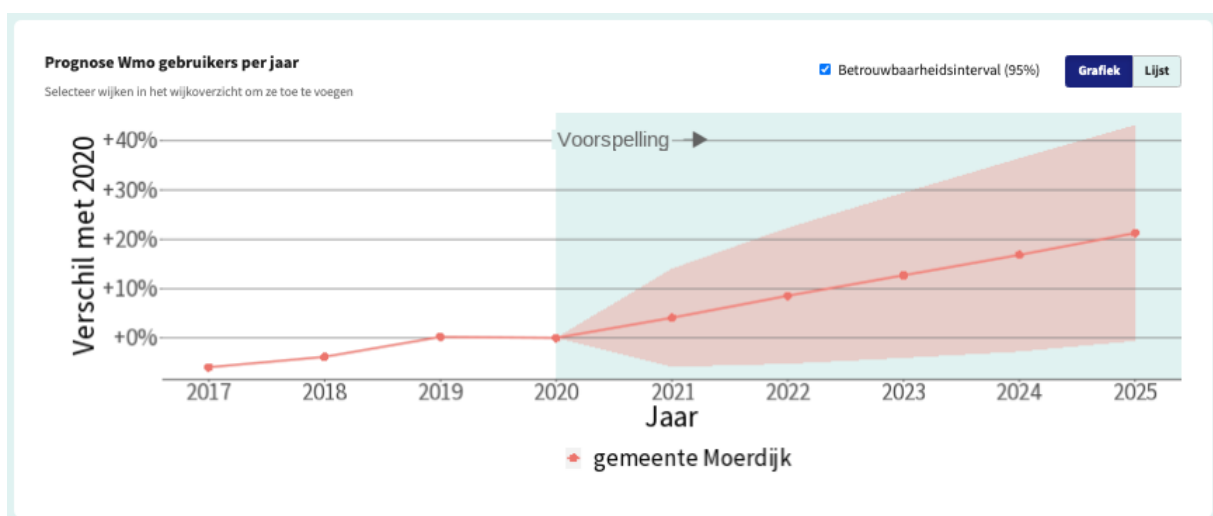
Het feit dat een voorspelmodel altijd een bepaalde mate van onzekerheid met zich meebrengt, maakt dat er altijd een kans bestaat dat de voorspelling onjuist is. Alle respondenten zijn zich bewust van het feit dat voorspelmodellen niet dé waarheid aangeven (R1-10). Dit geeft het voorspelmodel aan door middel van onzekerheidsmarges (R3, 8). Desalniettemin is wel sprake van vertrouwen in het model, zoals R2 stelt:

“Ik heb wel vertrouwen in dat het resultaat wat het model geeft goed is. Alleen ja, het blijft een glazen bol en daar proberen we natuurlijk zo dichtbij mogelijk te komen.”

Daarbij is het zo dat de onzekerheid van de voorspelling steeds groter wordt naarmate verder in de toekomst wordt gekeken (R1, 3). Dit belemmert echter voor R9 (business-analist) niet de bruikbaarheid van het model:

“Van de week zagen we ook weer dat een wijk echt wat betreft het aandeel Wmo-cliënten enorm groot zal zijn over vijf jaar. Ja, dan ligt dat misschien uiteindelijk vijf procentpunten hoger of lager. Maar het beeld blijft hetzelfde dat dat een wijk is die op dat punt aandacht vraagt.”

Echter staat een té grote onzekerheidsmarge de (verdere) doorwerking van het model wel degelijk in de weg, omdat de voorspelling in zo'n geval weinig informatief is over een mogelijke toe- of afname (R5 – data-analist). Ter illustratie laat Afbeelding 5 zien hoe een voorspelling eruitziet inclusief onzekerheidsmarges. Deze worden in het model betrouwbaarheidsintervallen genoemd. Wat ook de doorwerking beperkt, zijn de generieke categorieën waar het model een voorspelling voor geeft (R9, 10). Volgens R3 (senior beleidsmedewerker) is de keuze voor de huidige mate van detail desalniettemin een goede keuze geweest en is het uitbreiden naar een lager aggregatieniveau iets waar men geleidelijk naartoe kan werken. Dit komt namelijk de landelijke bruikbaarheid van het model ten goede, evenals het behouden van een acceptabel aggregatieniveau in termen van privacy (zodat personen niet herleidbaar zijn). Meerdere respondenten zouden de doorontwikkeling van het model graag gericht zien op de lokale toepasbaarheid door bijvoorbeeld de Wmo-voorzieningen nader te specificeren (R1, 3, 5, 9, 10).



Afbeelding 5 Voorspelling Wmo-gebruikers van Gemeente Moerdijk met betrouwbaarheidsinterval. Aangepast overgenomen uit *Wmo voorspelmodel* van VNG, n.d. (<https://wmo voorspelmodel.vng.nl/>). Copyright, n.d., VNG.

5.3.3.2 Procedurele output-legitimiteit

De manier waarop wordt omgegaan met de gegeven voorspelling wordt op verschillende aspecten beoordeeld. Het juist omgaan met de voorspelling vraagt om adequate kennis, expertise en capaciteit. Om als gemeente hierin te kunnen voorzien, is volgens respondenten een gestructureerde aanpak noodzakelijk.

Een belangrijke toegevoegde waarde van het voorspelmodel is dat de voorspelling informatie geeft waarover je het gesprek kunt aangaan (R7 – strategisch adviseur). De voorwaarde om dit juist te laten verlopen is dat de gebruiker adequate kennis en expertise bezit. Hiermee wordt bedoeld dat gebruikers de voorspelling kunnen interpreteren met een bepaald bewustzijn over de beperkingen van het model (R1-3, 7, 9). Dit vraagt om het verantwoord omgaan met de gegeven voorspelling (R9 – business-analist). Zo dient altijd te worden gecommuniceerd over de onzekerheidsmarges en de beperkte betrouwbaarheid naarmate het model verder in de toekomst kijkt (R3, 4,). Dit wijst op het aspect dat nog altijd sprake moet zijn van een menselijke maat (R7). Externe omstandigheden kunnen namelijk van invloed zijn op het Wmo-gebruik, maar kunnen niet worden meegenomen in het voorspelmodel. Een voorbeeld is de COVID-19 pandemie, maar ook lokale ontwikkelingen zoals nieuwbouw (R1, 3, 4, 5, 7). Zo vertelt R7:

“Het zou zomaar kunnen dat het model het ene laat zien, maar dat we weten dat in een bepaalde wijk een ontwikkeling plaatsvindt, bijvoorbeeld nieuwbouw. Voor Wmo is zoiets als nieuwbouw super interessant, want in nieuwbouw kun je bijvoorbeeld rekening houden met trapliften, of een gewone lift, die je in oudere huizen minder snel verwacht. En dat weet dan weer een beleidsadviseur en die kan dan ingrijpen in het model.”

Het belang van het verzekeren van het juiste gebruik van het model, blijkt uit het feit dat R4 (onderzoeker) samen met andere data-experts een training heeft verzorgd. Tijdens deze training is aandacht gegeven aan de onzekerheid van de voorspellingen (R4). Dit verzekert het goed kunnen duiden van de voorspelling en het vermogen om die waar nodig te nuanceren. Die onzekerheidsmarges belemmeren echter de communicatie naar bepaalde eindgebruikers die liever exacte percentages horen, zoals gemeenteraadsfracties (R3 – senior beleidsmedewerker). Met eindgebruikers worden de actoren bedoeld waar respondenten hun interpretaties van de voorspellingen aan presenteren.

Meerdere respondenten geven aan dat een gestructureerde aanpak vanuit de gemeente een essentiële stap is om verdere doorwerking van het voorspelmodel te kunnen faciliteren (R3, 4, 7, 9, 10). Een structurele aanpak voorkomt dat aandacht voor het model wegebt (R3, 7). Voor R10 (beleidsadviseur) is structuur belangrijk omdat doorwerking van het voorspelmodel meerdere teams aangaat binnen het sociaal domein:

“We willen eigenlijk vaker met elkaar rond de tafel, dus we gaan nu in overleg om dat structureler in te richten en de behoefte van die collega’s vanuit mijn team. [...] Dus we zijn op zoek naar een overlegvorm waar het sociaal domein in brede zin vertegenwoordigd is.”

Op die manier zou het model onderdeel kunnen worden van de bestaande werkwijze en wordt het Wmo-voorspelmodel een van de tools die medewerkers tot hun beschikking hebben om beter te kunnen sturen op beleid of om tussentijdse aanpassingen te kunnen doen (R10).

5.3.4 Feedback-legitimiteit

Feedback-legitimiteit draait om de terugkoppeling van verbeterpunten, adviezen, op- en aanmerkingen omtrent (het omgaan met) het voorspelmodel. Dit is inhoudelijk legitiem als de genoemde feedback-punten qua inhoud juist zijn (Bekkers, 2017). Het is procedureel legitiem als het omgaan met deze feedback-punten van voldoende kwaliteit is.

5.3.4.1 Inhoudelijke feedback-legitimiteit

Over het algemeen heersen onder de respondenten dezelfde wensen omtrent de doorontwikkeling van het voorspelmodel, namelijk de lokale toepassing verbeteren (R3, 5, 9, 10). R4 (onderzoeker) is zich bewust van deze wens en begrijpt ook dat die behoefte er is, maar geeft ook aan dat bijvoorbeeld het doen van uitspraken op voorzieningenniveau niet rechtstreeks mogelijk is. Om een voorspelling te kunnen doen, moeten namelijk de aantallen (aantal gebruikers) groot genoeg zijn (R4). Zo zeggen de generieke categorieën iets over Hulp bij Huishouden, Ondersteuning Thuis en Hulpmiddelen & Diensten, maar bijvoorbeeld niet over hoeveel mensen kunnen worden verwacht met een scootmobiel. Anderzijds is ook vanuit de beleid-kant begrip over de beperkte doorontwikkelingsmogelijkheden, zoals blijkt uit de reactie van R3 (senior beleidsmedewerker):

“Ik wil in principe het liefst alle trends weten op zeg maar alle voorzieningen waarover ik rapporteer. [...] Maar dat kan niet, want ze hebben een heel aantal voorzieningen op grote hopen bijeen verzameld. En daar heb ik alle begrip voor op het moment dat [data-specialist] vanuit zijn expertise zegt van ja, maar dan krijg je veel te kleine aantallen.”

Wat gebruikers wel kunnen doen is een soort vertaalslag maken op basis van de inzichten die uit het voorspelmodel kunnen worden gehaald (R4). Hiervan is R10 (beleidsadviseur) zich bewust, maar geeft tevens aan die omschakeling naar de lokale situatie lastig te vinden door de algemeenheid van de categorieën. Anderzijds begrijpen respondenten ook dat te gedetailleerde categorieën ten koste kan gaan van de betrouwbaarheid van het model (R3, 5).

Een andere wens is om de kwaliteit van het model te testen, waarmee wordt bedoeld dat de gegeven voorspelling wordt vergeleken met de uiteindelijke feitelijke cijfers (R2, 7, 9). Hierover vertelt R4:

“Op het moment dat er nieuwe data komt, dan kan je je prognose iets bijstellen op basis van de nieuwe inzichten die met je data te zien is. Dus op die manier houden we de vinger aan de pols. Het is niet zo dat we nu zeggen van nou, hij is klaar en we kijken er niet meer naar.”

5.3.4.2 Procedurele feedback-legitimiteit

Over de manier waarop wordt omgegaan met feedback op het model of het gebruik, zien respondenten dat het belangrijk is om het model als dynamisch te zien; het model is nooit af en moet in ontwikkeling blijven om betrouwbaar te zijn (R1, 4). Wat hierbij van belang is, is het in gesprek blijven met betrokken collega's (R7, 9, 10). Vragen waarbij wordt stilgestaan, gaan over de keuzes die in het verleden zijn gemaakt; op basis van welke aannames zijn die gemaakt en zijn die nog geldig? Om uitgebreid te kunnen beoordelen of de omgang met gegeven feedback juist is, lijkt het model nog in een te vroeg stadium te zijn, gezien de relatief recente lancering van het landelijke model (R2 – business- en data-analist). De VNG heeft in de zomer 2022 een evaluatieonderzoek gepland, waarover R2 het volgende stelt:

“Dan krijgen we evaluatie van het model en ga eens kijken in hoeverre klopt het model nou met wat wij hadden voorspeld? [...] En dan zullen er heus wel aanpassingen op het model gaan komen. Maar hoe of wat? Ja, dat moet allemaal nog gebeuren. Het model

zit ontzettend goed in elkaar, maar dat staat wat dat betreft nog wel in de kinderschoenen.”

Ondanks dat het model nog in de kinderschoenen staat, geven respondenten aan dat het participeren in het project al een vooruitgang is in het data-gedreven werken door gemeenten (R1, 5, 6, 9, 10). Die toegevoegde waarde wordt door R6 (data-engineer) als volgt geformuleerd:

“Ook al wordt het nu niet veel gebruikt, het kennismaken met dit soort manieren van werken helpt uiteindelijk wel voor een verdere doorontwikkeling, hè? Het is een stapje in een proces en het is ook koudwatervrees bij de gemeenten. Als je dit ooit wil gaan gebruiken, zul je dit langzaam stapje voor stapje moeten opbouwen. Ik bedoel, ja, zo gaan die dingen wel een beetje.”

Niet alleen regionale samenwerking is hierin van belang, maar ook tussen verschillende domein en data-experts (R1, 9, 10). Hierin herkent R10 (beleidsadviseur) een wederzijds afhankelijke relatie om de mogelijkheden van het voorspelmodel optimaal te kunnen benutten.

5.3.4.3 Sub-conclusie

In de beoordeling van de inhoudelijke input-legitimiteit geven respondenten aan dat de voorspelkracht, data-kwaliteit en het doel van het voorspelmodel van belang zijn. Om de procedurele input-legitimiteit als juist te beoordelen vinden zij een gestructureerde samenwerking essentieel. Wat betreft de onderliggende algoritmen staat de inhoudelijke *throughput*-legitimiteit van het Wmo-voorspelmodel nauwelijks ter discussie, doordat respondenten vertrouwen op de expertise van collega's die over de onderliggende algoritmen gaan. Het vertrouwen op die expertise borgt voor de respondenten veelal ook het als juist beoordelen van de procedurele *throughput*-legitimiteit. Goede onderlinge communicatie tussen domein- en data-experts is hierin belangrijk.

Over de inhoudelijke output-legitimiteit van het model wordt benadrukt dat de voorspellingen niet dé waarheid aantonen, maar dat altijd sprake is van een bepaalde mate van onzekerheid. Wat een acceptabele mate van onzekerheid is, moet volgens respondenten goed worden bediscussieerd onder de betrokken actoren. De ietwat generieke categorieën waarover het voorspelmodel een voorspelling geeft, lijkt de perceptie van de inhoudelijke output-legitimiteit te verminderen. Om de procedurele output-legitimiteit te beoordelen, zijn kennis,

expertise en capaciteit veelvoorkomende aspecten die worden genoemd. Daarbij is het volgens respondenten ook van belang dat gemeenten een gestructureerde aanpak hanteren in het raadplegen van het Wmo-voorspelmodel.

Wanneer het gaat om inhoudelijke feedback-legitimiteit, noemen respondenten vanuit de kant van beleid veelal dezelfde wensen over de doorontwikkeling van het model. Tegelijkertijd beseffen respondenten dat die wensen in het landelijke voorspelmodel lastig te realiseren zijn wanneer de landelijke inzetbaarheid behouden moet worden. Over de procedurele feedback-legitimiteit benadrukken respondenten dat deze dynamisch dient te zijn door het gesprek aan te blijven gaan met betrokken collega's. Samenwerking tussen domein- en data-experts en tussen verschillende gemeenten wordt hierin van belang geacht. Niet alleen om het voorspelmodel door te ontwikkelen, maar ook om data-gedreven werken in het algemeen verder te brengen.

6. Conclusie en discussie

Allereerst zullen de conclusies op de empirische deelvragen worden gegeven, om vervolgens de hoofdvraag te beantwoorden. Wat volgt is een discussie van de bevindingen in het licht van het theoretisch kader uit Hoofdstuk 3.

6.1 Conclusie

Het doel van dit onderzoek is om kennis te vergroten en theorie te ontwikkelen over de wijze waarop percepties van legitimiteit van ambtenaren de doorwerking beïnvloeden van algoritmen in de beleidsvorming. In dit onderzoek is de definitie van algoritmen van Leeuw (2019) gevolgd: ‘een lijst van stappen/instructies om met een computerprogramma een probleem op te lossen’ (p. 4). In dit onderzoek ging het om de percepties van legitimiteit van het algoritmische model. Daarbij is de volgende definitie van legitimiteit gehanteerd: een algoritmisch model is legitiem – in het ontwerp, de gegeven uitkomst, het gebruik en herziening – wanneer het past binnen de normen- en wettelijke kaders vanuit de perceptie van de betrokken actoren. Legitimiteit kan worden beoordeeld in zowel procedurele als inhoudelijke termen (Wallner, 2008). Het algoritmische model dat onder de loep is genomen, is het Wmo-voorspelmodel en kan worden gecategoriseerd als een *machine learning*-model dat leert en voorspellingen maakt van grote hoeveelheden data door patroonherkenning (Leeuw, 2020).

Om het onderzoeksdoel te verwezenlijken is een kwalitatieve casestudy uitgevoerd naar het Wmo-voorspelmodel, waarbij het semigestructureerde interview de voornaamste onderzoeksmethode is geweest. De theorie die wordt ontwikkeld zal de vorm krijgen van theoretische mechanismen. In dit hoofdstuk zullen die mechanismen worden beschreven door antwoord te geven op de hoofdvraag:

“Op welke wijze beïnvloedt de perceptie van legitimiteit van ambtenaren de doorwerking van het Wmo-voorspelmodel in de beleidsvorming in Nederlandse gemeenten?”

Het antwoord op de hoofdvraag wordt afgeleid door eerst de drie deelvragen te beantwoorden:

1. Op welke wijze is sprake van doorwerking van het Wmo-voorspelmodel in de beleidsvorming in verschillende Nederlandse gemeenten?
2. Wat zijn de percepties van legitimiteit van ambtenaren over het Wmo-voorspelmodel in verschillende Nederlandse gemeenten?

3. Wat is de relatie tussen de percepties van legitimiteit van ambtenaren over het Wmo-voorspelmodel en de doorwerking ervan in beleidsvorming in verschillende Nederlandse gemeenten?

Op de vraag op welke wijze sprake is van doorwerking van het Wmo-voorspelmodel in de beleidsvorming is het antwoord dat dit per type doorwerking varieert. Van doorwerking is in dagelijkse beleidshandelingen weinig tot geen sprake, doordat het voorspelmodel niet is bedoeld voor dagelijkse raadpleging, maar eerder maandelijks of naar aanleiding van specifieke actiepunten. Hoewel ook dit nog beperkt is, is van doorwerking in beleidsdocumenten enkel sprake in gemeenten waar een eigen Wmo-voorspelmodel in gebruik is.

Over de percepties van legitimiteit van ambtenaren over het Wmo-voorspelmodel kan worden gesteld dat dit verschilt per aspect van het model. Om tot een positieve beoordeling te komen wat betreft de data-input is voldoende voorspelkracht, data-kwaliteit en het doel van het voorspelmodel van belang. Wanneer de input inhoudelijk op orde is, komt dat de kwaliteit van de uiteindelijke voorspelling ten goede. Of het doorlopen ontwerpproces van het model van voldoende kwaliteit is, staat nauwelijks ter discussie, omdat over het algemeen wordt vertrouwd op de gestructureerde samenwerking tussen collega's met de gepaste expertise. Dit onderlinge vertrouwen in het beoordelingsvermogen van collega's is in dit verband essentieel. De aspecten die het juist omgaan met de voorspelling verzekeren, zijn de aanwezigheid van voldoende kennis, expertise en capaciteit binnen gemeenten. Opvallend is dat men het erover eens is dat het model zorgvuldig en goed in elkaar is gezet, maar dat de uiteindelijke voorspellingen die het model geeft niet de gewenste mate van detail hebben. Met name de generieke categorieën waarover het model voorspelt, werken beperkend, hetgeen meteen een aanzet geeft tot het beantwoorden van de derde deelvraag.

De derde deelvraag verbindt de eerste en de tweede door de relatie tussen beiden te bezien. Door het goede vertrouwen in collega's dat zij het ontwerpproces juist en zorgvuldig hebben doorlopen, wordt ook het model zelf over het algemeen als inhoudelijk goed beoordeeld. Zoals hierboven is aangestipt, zijn het de generieke categorieën waarover de voorspellingen gaan, die de doorwerking van het model limiteren. Respondenten zouden graag zien dat de voorspelling meer toegespitst is op de lokale context. Dit botst enerzijds met het uitgangspunt van het landelijke Wmo-voorspelmodel, namelijk de landelijke inzetbaarheid. Anderzijds botst dit met wat überhaupt mogelijk is met zo'n algoritmisch model om een bepaalde mate van betrouwbaarheid en nauwkeurigheid te behouden.

Zodoende kan uit bovenstaande het antwoord op de hoofdvraag worden geformuleerd als zijnde dat percepties van legitimiteit een beperkte invloed hebben op de doorwerking van het Wmo-voorspelmodel in de beleidsvorming, waarbij een positieve beoordeling van legitimiteit de doorwerking ervan niet garandeert. De invloed is beperkt omdat naast legitimiteit ook andere aspecten een voorwaarde zijn voor doorwerking van het voorspelmodel, zoals bruikbaarheid. Zo wordt het Wmo-voorspelmodel als landelijk model over het algemeen als legitiem beoordeeld. Men ziet in dat de meer generieke voorspellingen de brede inzetbaarheid en betrouwbaarheid van het landelijke model borgt. Echter, houdt de beperkte lokale bruikbaarheid de doorwerking van het model in beleidsvorming tegen.

Op basis van deze conclusie kan een theoretisch mechanisme worden opgesteld die het verband tussen percepties van legitimiteit en doorwerking van een algoritmisch model in beleidsvorming beschrijft. Zoals hierboven is genoemd, is een overwegend positieve beoordeling van een algoritmisch model in termen van legitimiteit niet voldoende voor doorwerking in beleidsvorming. Die beperking zit met name in de output-legitimiteit. Deze is inhoudelijk legitiem als de voorspelling als juist wordt beoordeeld. Het begrip ‘juist’ dient hierin echter nader te worden gespecificeerd. In de context van voorspelmodellen is het namelijk zo dat men altijd te maken heeft met bepaalde onzekerheidsmarges. In die zin is het daarom raadzaam het begrip ‘correct’ te gebruiken in plaats van ‘juist’. Die mogelijke incorrectheid van een voorspelling is inherent aan voorspelmodellen en maakt het model – vanuit de perceptie van respondenten – niet minder legitiem. Wanneer het begrip ‘juist’ wordt bedoeld als zijnde juist voor het doel waarvoor respondenten de voorspelling zouden willen gebruiken, is dit een ander verhaal. Die inhoudelijke juistheid vermindert namelijk als deze wordt getoetst aan de hand van de lokale toepasbaarheid. Zo kan de perceptie van output-legitimiteit de bruikbaarheid en daarmee de doorwerking van een algoritmisch model beïnvloeden.

6.2 Discussie

Zoals is behandeld in Hoofdstuk 3 wordt in dit onderzoek de benadering van de informatie ecologie gevolgd. Vanuit deze lens zijn de bevindingen geanalyseerd binnen de context van *algorithmic governance* (AG). Informatie ecologie legt de nadruk op (politieke en/of sociale) afwegingen van de mens omtrent bijvoorbeeld het ontwerp van een model en de manieren waarop deze keuzes het gebruik en de doorwerking ervan in het openbaar bestuur beïnvloeden (Latzer & Festic, 2019; Moody & Bekkers, 2022). Toegespitst op het landelijke Wmo-

voorspelmodel is het de keuze om het model landelijk inzetbaar te maken die de bruikbaarheid en daarmee de doorwerking in grote mate beïnvloedt. Hierin is de verdeelde *agency*/autonomie tussen mens en technologie te herkennen (Latzer & Festic, 2019). In dit licht heeft namelijk enerzijds de mens een bepaalde mate van autoriteit behouden. Zo waren het de betrokken actoren die de indeling in de drie geaggregeerde voorspelcategorieën (Hulp bij Huishouden, Ondersteuning Thuis en Hulpmiddelen & Diensten) vaststelden. Dit maakte het doen van generieke voorspellingen en daarmee de landelijke inzetbaarheid mogelijk. Anderzijds heeft de gebruikte techniek sterk de mogelijkheid om op voorzieningenniveau voorspellingen te doen beïnvloedt.

Respondenten benadrukten met klem dat gebruikers van het voorspelmodel in staat moeten zijn om dit te interpreteren met alle mitsen en maren. Dit wijst op de mate waarin de notie van *humans in the loop* in deze casus van belang is, zoals ook De-Artega et al. (2020) en Selbst en Barocas (2018) opperden. Hieraan gerelateerd is dat het Wmo-voorspelmodel geen *black box* is, maar een (zij het enkel voor experts) navolgbaar *white box*-model. Hoewel vragen omtrent transparantie en begrijpelijkheid van belang blijven, lijken de risico's van complexe *black box*-modellen waar meerdere auteurs voor waarschuwden, zoals Henman (2020), Janssen en Kuk (2016), Leeuw (2020), Selbst en Barocas (2018) en Van den Braak en Choenni (2017), bij het Wmo-voorspelmodel minder van toepassing. Ook het vermeende voordeel van personalisatie is in deze casus niet van toepassing waar Levy et al. (2021) op wezen. Het Wmo-voorspelmodel is namelijk niet gericht op het voorspellen van individueel gedrag, maar op het voorspellen van geaggregeerde cijfers.

In lijn met wat Levy et al. (2021) stelden, heeft de keuze voor algoritmische modellen als tool belangrijke implicaties voor zowel de probleemdefiniëring als de geopperde oplossing. Uit de resultaten blijkt dat de toegevoegde waarde van het Wmo-voorspelmodel is om huidige beleids- en uitvoeringsprocessen te ondersteunen, door inzichtelijk te krijgen waar en wanneer in de toekomst welke zorg (ongeveer) nodig is. De keuze om te investeren in deze tool brengt de manier waarop gemeenten kijken naar de huidige maatschappelijke problemen omtrent de Wmo aan het licht, namelijk als een probleem dat draait om vraag en aanbod waar zij op willen anticiperen. Gemeenten willen de toekomstige zorgvraag inzichtelijk krijgen, om daar hun zorgaanbod op aan te passen. Deze casus is daarmee een voorbeeld van de notie van een data-gedreven *Anticipatory Government* (Maffei, Leoni & Villari, 2020). In plaats van blind sturen maakt een anticiperende overheid gebruik van vooruitziendheid (Fuerth, 2009). Een algoritmisch model – zoals het Wmo-voorspelmodel – kan worden gezien als een concrete tool voor vooruitziendheid wat data-gedreven beleidsvorming mogelijk maakt (Maffei et al., 2020).

Zo speelt het model tevens in op het optimistische idee over algoritmen, zoals gesteld door De-Artega et al. (2020), om prestaties van het openbaar bestuur te verbeteren door middel van preciezer voorspellingen. Met dit doel als uitgangspunt werd het voorspelmodel ontworpen. Janssen en Kuk (2016) stelden dat ook institutionele logica's, ideeën en overtuigingen van ontwerpers en beleidsmakers het ontwerp van het model beïnvloeden. In het kader van de ontwikkeling van het landelijke voorspelmodel is hierin duidelijk de landelijke bruikbaarheid leidend geweest, omdat dat het gewenste resultaat was waarnaartoe werd gewerkt.

Om data-gedreven te kunnen werken is zowel kennis van de inhoud als van data vereist. In de praktijk is deze kennis vaak verdeeld over domein- en data-experts. In deze taakverdeling is een van de principes van een Weberiaans bureaucratisch bestuurlijk systeem te herkennen, namelijk dat specialisatie specificereert wat men van wie kan verwachten (Mahfooz, 2017). Een voordeel hiervan is dat het werknemers ondersteunt in het ontwikkelen van expertise wat hun prestatie verbetert. In de casus van dit onderzoek blijkt dat door deze taakverdeling, samenwerking tussen domein- en data-experts van groot belang is. Hier zit een mogelijke beperking van een Weberiaans bureaucratisch bestuur, namelijk dat werknemers niet verder denken en doen dan hun vastgestelde taken (Mahfooz, 2017). Door die specialisatie vinden respondenten het namelijk niet noodzakelijk dat alle betrokken ambtenaren de inhoudelijke kwaliteit van bijvoorbeeld de onderliggende algoritmen kunnen beoordelen. Zo wordt over het algemeen de legitimiteit wat betreft de (data-)input en *throughput* (onderliggende algoritmen) voor lief genomen door het vertrouwen in het beoordelingsvermogen van collega's.

Waar wel op wordt gehamerd is het kunnen interpreteren van de voorspelling door gebruikers van het model. Dit wijst op het belang van de procedurele output-legitimiteit, omdat dit gaat over het omgaan met de gegeven voorspelling. Respondenten geven aan dat bij gebruik van het Wmo-voorspelmodel bewustzijn van alle mitsen en maren van belang is, zoals het gegeven dat een voorspelling niet dé waarheid vertoont. Dit heeft tevens te maken met inhoudelijke output-legitimiteit. Jagers et al. (2016) stelden namelijk dat de output-legitimiteit toeneemt als een voorspelling gunstig is voor de betrokkenen. In andere woorden kan de output als legitiem worden beoordeeld wanneer dit bijdraagt aan het effectief oplossen van maatschappelijke problemen (Eshuis & Edwards, 2013). Dit haakt in op het theoretisch mechanisme dat in de vorige paragraaf is beschreven, namelijk de nadere specificatie van het begrip 'juist' in de zin van het beoordelen van de voorspelling als inhoudelijk juist.

Deze casus kan die notie van output-legitimiteit op twee manieren aanvullen. Enerzijds hoeft een gegeven voorspelling namelijk niet 100% juist (correct) te zijn voor respondenten om als legitiem te beoordelen. Inherent aan een voorspelmodel is namelijk dat sprake is van een

bepaalde (acceptabele) mate van onzekerheid. Het bewustzijn over deze onzekerheid is essentieel voor het juist omgaan met de voorspelling. Zo wordt de beperkte inhoudelijke legitimiteit volgens respondenten gecompenseerd door een hoge mate van procedurele legitimiteit. Anderzijds, dient de voorspelling wel juist te zijn voor het beoogde doel van respondenten. Het zijn namelijk de geaggregeerde categorieën waarover het model een voorspelling geeft, die uiteindelijk de output-legitimiteit beperken. Output-legitimiteit blijkt hierdoor wel degelijk een belangrijke rol te spelen wanneer de definitie van Eshuis en Edwards (2013) wordt gevolgd, namelijk als het vermogen om maatschappelijke problemen aan te pakken. De categorieën van het Wmo-voorspelmodel zijn dermate generiek dat het landelijke model niet als voldoende lokaal toepasbaar wordt geacht, waardoor de doorwerking in beleidsvorming beperkt is.

6.3 Vervolgonderzoek

Omdat in dit onderzoek enkel een theoretisch mechanisme is opgesteld, is vervolgonderzoek noodzakelijk om een verdere theorie te kunnen construeren en om de validiteit en consistentie van dit onderzoek te toetsen (Van Thiel, 2014). Met name de concepten van legitimiteit binnen de context van algoritmische modellen dienen nader onderzocht te worden in de bruikbaarheid en validiteit ervan. Deze zijn namelijk naar eigen inzicht toegepast op de verschillende deelaspecten van het Wmo-voorspelmodel en de ontwikkeling ervan.

Ander relevant vervolgonderzoek zou zijn om een vergelijkende studie te doen naar *white box*- en *black box*-modellen. Zoals is benoemd in de vorige paragraaf, bleek dat in deze casus minder sprake is van de risico's waar auteurs voor waarschuwen wanneer zij het hebben over algoritmische modellen. Het is interessant om te bekijken welke voor- en nadelen gebruikers van een complexer voorspelmodel in een vergelijkbare context ervaren.

7. Reflectie

Dit hoofdstuk bediscussieert de methodologische sterke en zwakke kanten van dit onderzoek.

Ten eerste is de relatief korte periode waarin dit onderzoek is uitgevoerd een beperking geweest. Het doel van deze casestudy was om een zo holistisch mogelijk begrip van de casus te verkrijgen (Van Thiel, 2014). Idealiter stopt de onderzoeker met dataverzamelen wanneer verzadiging optreedt. Dit was in dit onderzoek niet haalbaar, wat maakte dat een aantal pragmatische keuzes zijn gemaakt om genoeg te nemen met tien respondenten. Om dezelfde reden zijn ook een beperkt aantal documenten geanalyseerd. In een langduriger onderzoek was dit niet nodig geweest. Dit vermindert de externe validiteit, omdat niet met zekerheid is te stellen of de bevindingen representatief zijn voor alle gemeenten die met het Wmo-voorspelmodel werken. Kwalitatief onderzoek is overigens over het algemeen minder extern valide in vergelijking met kwantitatief onderzoek (Van Thiel, 2014).

Een ander minder sterk punt gaat om de interne validiteit, vanwege de lastig empirisch observeerbare variabelen van dit onderzoek (Van Thiel, 2014). De onafhankelijke variabele ‘percepties van legitimiteit’ is bij de operationalisering opgedeeld in vier deelconcepten met meerdere dimensies. Deze analytische onderscheiding was soms lastig te behouden tijdens metingen in de empirie. Dit maakte ook het correct coderen en interpreteren van de data een uitdaging. Door continue het operationaliseringschema bij het coderen ernaast te houden, is dit probleem geminimaliseerd.

Ook de gebruikte onderzoeksmethode van interviews is enigszins beperkend geweest. Hoewel dit een geschikte methode is om percepties van respondenten te meten, maakte het semigestructureerde karakter dat regelmatig is afgeweken van de vooraf opgestelde vragenlijst. Zo zijn een aantal vragen overgeslagen of juist toegevoegd door het verloop van het gesprek. Feitelijk is hierdoor meer gebruikgemaakt van een thematische topiclijst dan van de gedetailleerde vragenlijst zoals opgesteld in Bijlage 3. Dit heeft mogelijk de vergelijkbaarheid, interne validiteit en betrouwbaarheid van de theoretische constructen verminderd (Van Thiel, 2014). Aan de andere kant bood een thematische topiclijst de mogelijkheid om meer het verloop van het gesprek aan te houden en de beleving van de respondenten te volgen. Hierdoor werd de validiteit weer verhoogd, omdat de metingen goed bij de empirische realiteit aansloten.

Wat de kwaliteit van het onderzoek heeft bevorderd, was de ter voorbereiding uitgevoerde onderzoeksmethode van participerende observatie. Dit hielp bij een beter begrip van de ervaringen van respondenten, omdat de Leergang huidige uitdagingen en kansen van algoritmen in het openbaar bestuur behandelde. Dit maakte verdere diepgang mogelijk, wat

bijdroeg aan een beter beeld van de casus. Door een reflexieve houding en het vooraf opstellen van een niet-sturende vragenlijst is gepoogd om sturend vragen te voorkomen.

Wat betreft het theoretisch kader kan worden gesteld dat deze voor het empirische gedeelte van het onderzoek bruikbaar is gebleken. De lens van de informatie ecologie binnen de *algorithmic governance* heeft de data-analyse alert gemaakt op de meest relevante aspecten van de casus. Hoewel niet alle theoretische noties even relevant waren voor deze specifieke casus, hebben deze inzichten geleid tot een beter begrip ervan. Niet-toepasselijkheid van bepaalde theoretische noties uit het theoretisch kader betekent namelijk niet dat deze zinloos waren bij het verzamelen en analyseren van de data. Veel bestaande literatuur blijkt gericht te zijn op complexe algoritmische modellen, terwijl in het openbaar bestuur nog relatief weinig sprake is van gebruik van zulke complexe modellen. Hierdoor waren van de risico's waar auteurs het veelal over hadden niet van toepassing op de casus van dit onderzoek.

8. Aanbevelingen

Op 20 januari 2022 is het Wmo-voorspelmodel landelijk gelanceerd door de VNG voor gebruik door alle Nederlandse gemeenten. Na die lancering was het aan gemeenten zelf om aan de slag te gaan en aan te geven wat wel en niet goed gaat. Nu is het ruim een halfjaar later en hebben verschillende gemeenten op hun eigen manier hieraan gehoor gegeven. De aanbevelingen van dit onderzoek zijn enerzijds gericht aan gemeenten en aan de VNG anderzijds. Op voorhand dient te worden gesteld dat dit slechts één van de manieren is om vervolgstappen te ondernemen.

8.1 Aan gemeenten

Gemeenten hebben dankzij betrokkenheid bij (de ontwikkeling van) het Wmo-voorspelmodel (extra) ervaring op kunnen doen in data-gedreven werken. Deze opgedane ervaring werd door respondenten als waardevol ervaren. Indien gemeenten het data-gedreven werken (verder) willen doorzetten in de organisatie, is het aan te bevelen een data-gedreven werkwijze te creëren die onderdeel wordt van de structurele gang van zaken. Net als het *whitepaper* van DigiTime (2021) stelt, blijkt ook uit dit onderzoek dat het niet een kwestie van onwil is, maar een kwestie van kunnen en doen. Om de stap te zetten naar het kunnen en doen, is het advies om een permanent team aan te stellen dat zich bezighoudt met het volbrengen van data-gedreven projecten om deze uiteindelijk te consolideren in gemeentelijke werkwijzen. Vanwege grote verschillen in capaciteit en expertise tussen kleine, middelgrote en grote gemeenten, is het wenselijk om dit team op regionaal niveau op te stellen. Op deze manier wordt voorkomen dat data-gedreven projecten enkel met het juiste momentum en alleen door grote gemeenten volbracht kunnen worden. Het is belangrijk om het data-gedreven werken in de basis op orde te hebben om de potentie van het Wmo-voorspelmodel en andere toekomstige modellen te kunnen benutten, zoals in de beleidsvorming. Een goede basis bevordert namelijk kwalitatief goede data, wat essentieel is als gemeenten data-gedreven willen werken.

Om de legitimiteit van het Wmo-voorspelmodel te borgen, is het aan te raden om het doel van het model te herzien met het oog op de behoeftes en doelstellingen van de gemeente. Gemeenten kunnen voor zichzelf nagaan welk probleem dit model dient op te lossen en of het huidige model geschikt hiervoor is. Op basis hiervan kan worden beoordeeld of doorontwikkeling van dit model of ontwikkeling van een ander model logischer is. Wanneer het model en de doelstellingen van gemeenten goed op elkaar aansluiten, kan worden gekeken naar de mogelijkheden voor doorwerking in beleidsvorming.

8.2 Aan de VNG

Niet alleen binnen gemeenten, zoals hierboven beschreven, kan een structurele aanpak goed zijn. Ook tussen gemeenten zou een structureel samenwerkingsverband in data-gedreven werken potentie hebben. De ene gemeente is namelijk verder in het proces van data-gedreven werken dan de ander. Hierin kan de VNG een leidende rol spelen door de juiste teams per gemeente aan elkaar te koppelen, bijvoorbeeld op regionaal niveau. Ook hiervoor geldt dat dit een structurele plek dient te krijgen waarbij deze regionale teams regelmatig bij elkaar komen (bijvoorbeeld maandelijks of tweemaandelijks). Zo wordt de expertise en capaciteit van grotere gemeenten toegankelijk gemaakt voor kleinere gemeenten en is het mogelijk om te leren van elkaars ervaringen en uitdagingen. Om dit vorm te kunnen geven, is het aan te raden om als VNG allereerst met geïnteresseerde gemeenten apart in gesprek te gaan over wat zij uit zo'n regionale samenwerking zouden willen halen. Vervolgens is het aan de VNG om verschillende samenwerkingsverbanden tussen gemeenten op te zetten, met de behoeften en wensen van iedere gemeente in ogenschouw.

Evenals voor het Wmo-voorspelmodel specifiek is het voor de doorontwikkeling relevant om per gemeente in beeld te krijgen hoe de gemeente ervoor staat. Hoewel het Wmo-voorspelmodel voor landelijk gebruik is bedoeld, is uit dit onderzoek namelijk gebleken dat doorwerking van het model meer maatwerk vereist. Vragen die hierbij kunnen worden gesteld zijn: waar wil de gemeente het model het liefst voor gebruiken? Wat is daarvoor nodig volgens de betrokken domein- én data-experts? Welke modellen en/of dashboards gebruikt of ontwikkelt de gemeente zelf al en hoe verhouden die zich tot elkaar? Gezien de behoefte vanuit verschillende gemeenten om het model lokaal beter te kunnen toepassen, is het aan te raden om hieraan in het evaluatieonderzoek voldoende aandacht te besteden. Een overweging kan zijn om de landelijke inzetbaarheid van het model ietwat los te laten en te investeren in meer op maat gemaakte modellen. Men kan denken aan voorspelmodellen die gericht zijn op bepaalde typen gemeenten, zoals klein, middelgroot en groot, of andere overeenkomstige kenmerken. Het landelijke voorspelmodel kan dienen als instapmodel om gemeenten kennis te laten maken met dergelijke modellen en bijbehorende mogelijkheden.

Bibliografie

- Airaksinen, J., Härkönen, H., & Haveri, A. (2013). Perceptions of Legitimacy in Nordic Regional Development Networks. *Public Organization Review*, 14(4), 457-476.
- Algemene Rekenkamer. (2021). *Aandacht voor algoritmen*. Geraadpleegd op 10 februari 2022, van <https://www.rekenkamer.nl/publicaties/rapporten/2021/01/26/aandacht-voor-algoritmen>.
- Babbie, E. (2016). *The Practice of Social Research* (Fourteenth edition). Boston: Cengage Learning.
- Bekkers, V. (2017). *Beleid in beweging: Achtergronden, benaderingen, fasen en aspecten van beleid in de publieke sector* (3^e druk). Den Haag: Boom bestuurskunde.
- Boeije, H. (2010). *Analysis in Qualitative Research*. Londen: SAGE Publications Ltd.
- Bowen, S., & Zwi, A.B. (2005). Pathways to “Evidence-Informed” Policy and Practice: A Framework for Action. *PLoS medicine*, 2(7), e166.
- Craglia, M., Hradec, J., & Troussard, X. (2020). The Big Data and Artificial Intelligence: Opportunities and Challenges to Modernise the Policy Cycle. *Science for Policy Handbook*, 96-103.
- De-Arteaga, M., Fogliato, R., & Chouldechova, A. (2020). A Case for Humans-in-the-Loop: Decisions in the Presence of Erroneous Algorithmic Scores. *Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems*, 1-12.
- De Fine Licht, J. (2011). Do We Really Want to Know? The Potentially Negative Effect of Transparency in Decision Making on Perceived Legitimacy. *Scandinavian Political Studies*, 34(3), 183-201.
- De Jong, F. (2019). Haags Wmo-voorspelmodel krijgt landelijk vorm. Geraadpleegd op 17 maart 2022, van <https://vng.nl/sites/default/files/2021-06/haags-wmo-voorspelmodel-krijgt-landelijk-vorm-8-10-2019.pdf>.
- De Lange, M.A. (1995). *Besluitvorming rond strategisch ruimtelijk beleid: verkenning van doorwerking als beleidswetenschappelijk begrip* (Proefschrift). Amsterdam: Thesis Publishers.
- DigiTime. (2021). *Whitepaper DigiTime: analyse naar de digitale transformatie binnen overheidsorganisaties*. Geraadpleegd op 1 april 2022, van <https://www.linkedin.com/feed/update/urn:li:activity:6817842182775468032/>.

- Eshuis, J., & Edwards, A. (2013). Branding the City: The Democratic Legitimacy of a New Mode of Governance. *Urban Studies*, 50(5), 1066-1082.
- Fuerth, L. S. (2009). Foresight and anticipatory governance. *Foresight*, 11(4), 14-32.
- Gemeente Den Haag. (2020). *Datagedreven werken voor de stad. Datastrategie 2020-2022*. Geraadpleegd op 3 mei 2022, van https://denhaag.raadsinformatie.nl/document/8647666/1/RIS305091_Bijlage_2_Etalageprojecten_Datastrategie.
- Gemeente Den Haag. (2021). Algoritmeregister Den Haag. Geraadpleegd op 17 maart, 2022, van <https://www.dataplatform.nl/#/data/f58f2b0e-8d93-480c-b6f7-da95ed7bbe18>.
- Gritsenko, D., & Wood, M. (2022). Algorithmic governance: A modes of governance approach. *Regulation & Governance*, 16(1), 45-62.
- Head, B.W. (2015). Toward More “Evidence-Informed” Policy Making? *Public Administration Review*, 76(3), 472-484.
- Henman, P. (2020). Improving public services using artificial intelligence: possibilities, pitfalls, governance. *Asia Pacific Journal of Public Administration*, 42(4), 209-221.
- Höchtel, J., Parycek, P., & Schöllhammer, R. (2016). Big data in the policy cycle: Policy decision making in the digital era. *Journal of Organizational Computing and Electronic Commerce*, 26(1-2), 147-169.
- iBestuur. (2022, 25 januari). *Haags Wmo-voorspelmodel gaat landelijk*. Geraadpleegd op 17 maart 2022, van <https://ibestuur.nl/partner-vng-realisatie/haags-wmo-voorspelmodel-gaat-landelijk>.
- Issar, S., & Aneesh, A. (2021). What is algorithmic governance? *Sociology Compass*, 1-14.
- Jagers, S.C., Matti, S., & Nordblom, K. (2016). How Policy Legitimacy Affects Policy Support Throughout the Policy Cycle. Geraadpleegd op 10 februari, 2022, van https://gupea.ub.gu.se/bitstream/handle/2077/50725/gupea_2077_50725_1.pdf?sequence=1&isAllowed=y.
- Janssen, M. & Kuk, G. (2016). The challenges and limits of big data algorithms in technocratic governance. *Government Information Quarterly*, 33(3), 371-377.
- Kitchin, R. (2017) Thinking critically about and researching algorithms. *Information, Communication & Society*, 20(1), 14-29.
- Kolkman, D. (2020). The usefulness of algorithmic models in policy making. *Government Information Quarterly*, 37(3), 101488.
- Latzer, M., & Festic, N. (2019). A guideline for understanding and measuring algorithmic governance in everyday life. *Internet Policy Review*, 8(2), 1-19.

- Levy, K., Chasalow, K.E., & Riley, S. (2021). Algorithms and decision-making in the public sector. *Annual Review of Law and Social Science*, 17, 309-334.
- Leeuw, F.L. (2019). Evaluatieonderzoek, Big Data en Artificiële Intelligentie: een verkenning. *Beleidsonderzoek Online*.
- Leeuw, F.L. (2020). Program evaluation B: evaluation, big data, and artificial intelligence: two sides of one coin. In E. Vigoda-Gadot & R. Dana (Red.), *Handbook of Research Methods in Public Administration, Management and Policy* (pp. 227-297). Cheltenham: Edward Elgar Publishing.
- Maffei, S., Leoni, F., & Villari, B. (2020). Data-driven anticipatory governance. Emerging scenarios in data for policy practices. *Policy Design and Practice*, 3(2), 123-134.
- Mahfooz, A. (2017). The Theory of Bureaucracy of Max Weber, Merits and Demerits.
- Merad, M., & Trump, B.D. (2018). The legitimacy principle within French risk public policy: A reflective contribution to policy analytics. *Science of the Total Environment*, 645, 1309-1322.
- Moody, R. (2021, 13 december). Data-gedreven werken, In *ICTU 20 jaar* [Podcast]. Geraadpleegd op 26 januari 2022, van <https://www.ictu.nl/publicaties/podcast-datagedreven-werken>.
- Moody, R., & Bekkers, V. (2022). *Big data and public policy*. Palgrave, London (in print).
- Movisie (2013, 15 januari). *De drie decentralisaties in het sociale domein: Een overzicht in vogelvlucht*. Geraadpleegd op 12 juli 2022, van <https://www.movisie.nl/artikel/drie-decentralisaties-sociale-domein>.
- Muller, C. (2021, 30 November). Point of Action for Governmental AI Strategies. In B. Klievink (Voorzitter), *Better Algorithms for Better Policies*. Congres georganiseerd door STEPSS, online.
- Neuman, W.L. (2014). *Understanding Research* (Compiled by FSW Methoden en Statistiek Universiteit Utrecht). Harlow: Pearson Education Limited.
- Nilsson, M., Jordan, A., Turnpenny, J., Hertin, J., Nykvist, B., & Russel, D. (2008). The Use and Non-Use of Policy Appraisal Tools in Public Policy Making: An Analysis of Three European Countries and the European Union. *Policy Sciences*, 41(4), 335-355.
- Pencheva, I., Esteve, M., & Mikhaylov, S.J. (2020). Big Data and AI – a transformational shift for government: So, what next for research? *Public Policy and Administration*, 35(1), 34-44.
- Poel, M., Meyer, E.T., & Schroeder, R. (2018). Big Data for Policymaking: Great Expectations, but with Limited Progress? *Policy and Internet*, 10(3), 347-367.

- Punte, F. (2021, 6 juli). *Koplopergemeenten werken aan ontwikkeling Wmo voorspelmodel*. Geraadpleegd op 17 maart, van <https://kennisnetwerkdata.pleio.nl/groups/view/647a9dd1-3dc8-481c-abd2-86b631175ae7/wmo-voorspelmodel/blog/view/40526d37-8121-4f21-8031-b6942d0b2da0/koplopergemeenten-werken-aan-ontwikkeling-wmo-voorspelmodel>.
- Rich, R.F. (1997). Measuring Knowledge Utilization: Processes and Outcomes. *Knowledge and Policy: The International Journal of Knowledge Transfer and Utilization*, 10(3), 11-24.
- Rijksoverheid. (n.d.). *Wet maatschappelijke ondersteuning (Wmo)*. Geraadpleegd op 26 april, 2022, van <https://www.rijksoverheid.nl/onderwerpen/zorg-en-ondersteuning-thuis/wmo-2015>.
- Scharpf, F.W. (2009). Legitimacy in the multilevel European polity. *European Political Science Review*, 1(2), 173-204.
- Selbst, A.D., & Barocas, S. (2018). The Intuitive Appeal of Explainable Machines. *Fordham Law Review*, 87(3), 1085-1139.
- Steunenberg, B. (2018). Adaptieve beleidsontwikkeling: zoeken naar nieuwe vormen van beleidsanalyse voor de digitale overheid. *Beleidsonderzoek Online*.
- Van Buuren, A., Driessen, P., Teisman, G., & Van Rijswijk, M. (2014). Toward legitimate governance strategies for climate adaption in the Netherlands: combining insights from a legal, planning, and network perspective. *Regional environmental change*, 14(3), 1021-1033.
- Van Buuren, A., Klijn, E., & Edelenbos, J. (2012). Democratic Legitimacy of New Forms of Water Management in the Netherlands. *International Journal of Water Resources Development*, 28(4), 629-645.
- Van den Braak, S.W., & Choenni, R. (2017). *Het gebruik van datagedreven analysemethoden in de (beleids)praktijk: kansen, uitdagingen en handreikingen*. Geraadpleegd op 9 februari, 2022, van https://repository.wodc.nl/bitstream/handle/20.500.12832/928/Mem2017-3_Volledige_tekst_tcm28-301761.pdf?sequence=1&isAllowed=y.
- Van den Braak, S., & Choenni, S. (2019). Voorspellen met big-datamodellen. Over de valkuilen voor beleidsmakers. *Justitiële Verkenningen: De Toekomst Verkennen en Voorspellen*, 45(4), 21-38.
- Van der Voort, H.A.J., Klievink, A., Arnaboldi, M., & Meijer, A.J. (2019). Rationality and politics of algorithms. Will the promise of big data survive the dynamics of public decision making? *Government Information Quarterly*, 36(1), 27-38.

- Van Thiel, S. (2014). *Research Methods in Public Administration and Public Management: An Introduction*. London & New York: Routledge.
- Van Veenstra, A.F., & Kottering, B. (2017, September). Data-Driven Policy Making: The Policy Lab Approach. In *International conference on electronic participation* (pp. 100-111). Springer, Cham.
- Verhulst, S.G., Engin, Z., & Crowcroft, J. (2019). Data & Policy: A new venue to study and explore policy-data interaction. *Data & Policy*, 1, 1-5.
- VNG. (n.d.-1) *Wmo-voorspelmodel*. Geraadpleegd op 6 april, 2022, van <https://vng.nl/projecten/wmo-voorspelmodel>.
- VNG. (n.d.-2) *Wmo voorspelmodel*. Geraadpleegd op 17 juni, 2022, van <https://wmo-voorspelmodel.vng.nl/>
- VNG Realisatie. (2021). *Wmo voorspelmodel. Handreiking*. Geraadpleegd op 29 april, 2022, van <https://wmo-voorspelmodel.vng.nl/Handreiking%20Wmo%20voorspelmodel%20december%202021%20PDF.pdf>.
- Wallner, J. (2008). Legitimacy and Public Policy: Seeing Beyond Effectiveness, Efficiency, and Performance. *The Policy Studies Journal*, 36(3), 421-443.
- Wolfsen, A. (2021, 10 December). *Privacyblog Aleid Wolfsen: Als algoritmen discrimineren....* Geraadpleegd op 13 februari 2022, van <https://www.autoriteitpersoonsgegevens.nl/nl/nieuws/privacyblog-aleid-wolfsen-als-algoritmen-discrimineren>.
- WRR. (2021). Opgave AI: De nieuwe systeemtechnologie. Geraadpleegd op 9 maart 2022, van <https://www.wrr.nl/publicaties/rapporten/2021/11/11/opgave-ai-de-nieuwe-systeemtechnologie>.

Bijlage 1 Observatieschema Participerende Observatie

Datum	
Tijd	
Locatie	
Participant	
Onderwerp	

Concept	Indices	Indicatoren
Algoritmen	Simpele statistische analyses	Een algoritmisch model maakt statistische berekeningen middels bijvoorbeeld beslisbomen
	<i>Machine learning</i>	Een algoritmisch model leert en maakt voorspellingen van grote hoeveelheden data door patroonherkenning
	Artificiële intelligentie	Een algoritmisch model bepaalt autonoom zijn eigen interne regels om tot een uitkomst te komen

Concept	Indices	Indicatoren
Legitimiteit	Input-legitimiteit	Inhoudelijk
		Procedureel
	<i>Throughput</i> -legitimiteit	Inhoudelijk
		Procedureel
	Output-legitimiteit	Inhoudelijk
		Procedureel
	Feedback-legitimiteit	Inhoudelijk
		Procedureel

Bijlage 2 Respondentenlijst

Respondentnummer	Functie	Datum	Tijd	Locatie
1	Beleidsadviseur	29-04-2022	13:00-14:00	Zoom
2	Business en data analist OCW	04-05-2022	10:00-11:00	Teams
3	Senior beleidsmedewerker	10-05-2022	14:00-15:00	Teams
4	Onderzoeker	11-05-2022	15:00-16:00	Teams
5	Data-analist en coördinator vakgroep Kennis en Verkenning	12-05-2022	13:00-14:00	Teams
6	Data engineer	12-05-2022	13:00-14:00	Teams
7	Strategisch adviseur	23-05-2022	14:00-15:00	Teams
8	Data-scientist	23-05-2022	14:00-15:00	Teams
9	Business analist	03-06-2022	09:00-10:00	Teams
10	Beleidsadviseur	03-06-2022	09:00-10:00	Teams

Bijlage 3 Interviewprotocol

Vragenlijst:

Datum	
Tijd	
Locatie	
Respondent	

Opening:

- Waardering deelname
- Onderwerp onderzoek
- Onderwerp interview
- Interesse onderwerp
- Randvoorwaarden: opnemen, anoniem, transcriptie opsturen + mogelijkheid om erop te reageren

- Functie binnen de gemeente & betrokkenheid bij het Wmo-voorspelmodel
- Het gebruik van het Wmo-voorspelmodel door gemeente en uzelf:
 - o Functie: waarvoor wordt het gebruikt en/of waarvoor zou het kunnen worden gebruikt?
 - (Hoe) onderdeel van uw werk?
 - Frequentie: dagelijks, wekelijks, etc.?
 - Expliciet in beleidsdocumenten?
 - Wie zijn daarbij betrokken?
 - Toegevoegde waarde?
 - o Interesse: waar komt de interesse voor het model vandaan?
- Het model zelf:
 - o In hoeverre bent u bekend met hoe het model is ontworpen?
 - De data die als input hebben gediend?
 - Het algoritme om tot een output te komen?
 - o In hoeverre beoordeelt u de data die gebruikt zijn om het model te ontwerpen als inhoudelijk juist – als in: dient het de juiste normen en waarden?
 - o Wat kunt u vertellen over het doorlopen proces waarbij gekozen is om bepaalde data wel/niet mee te nemen?

- En wat betreft het algoritme? In hoeverre beoordeelt u de juistheid daarvan?
- Hoe is het besluit tot het gebruik van dit algoritme gekomen?
 - Idem: Vindt u dit proces juist doorlopen?
- In hoeverre vindt u dit van belang om het model te kunnen gebruiken?
 - Zowel in termen van de data-input als het algoritme?
- (Omgaan met) feedback voor (het gebruik van) het voorspelmodel:
 - Hoe wordt erop toegezien dat het model juist werkt/ blijft werken?
 - In termen van de data-input, het algoritme, het resultaat dat het model geeft, en hoe er met dat resultaat wordt omgegaan
 - Wat vindt u hiervan?
 - Is hiervoor een formele procedure?
 - In hoeverre vindt u dit belangrijk om het model te kunnen gebruiken?
- (Omgaan met) output van het voorspelmodel:
 - Kunt u me iets vertellen over hoe de output aan uiteindelijke gebruikers wordt gepresenteerd?
 - Hoe wordt er met die output omgegaan?
 - Bestaat hiervoor een formele procedure om te volgen?
 - Is er iemand die daar nog uitleg, verklaring, of interpretatie over/van geeft?
 - Is dat nodig volgens u?
 - In hoeverre beoordeelt u de juistheid van het resultaat dat het model geeft?
 - In hoeverre vindt u dit belangrijk om het model te kunnen gebruiken?

Bijlage 4 Documentenlijst

Documentnummer	Afkomstig van
1	Janneke Lummen
2	R3
3	R4
4	R4

Bijlage 5 Documentanalyse-protocol

Concept	Indices	Indicatoren
Algoritmen	Simpele statistische analyses	Een algoritmisch model maakt statistische berekeningen middels bijvoorbeeld beslisbomen
	<i>Machine learning</i>	Een algoritmisch model leert en maakt voorspellingen van grote hoeveelheden data door patroonherkenning
	Artificiële intelligentie	Een algoritmisch model bepaalt autonoom zijn eigen interne regels om tot een uitkomst te komen

Concept	Indicatoren	Items
Doorwerking	Doorwerking in beleidsdocumenten	Is er sprake van concrete uitwerking, invulling, vertaling en/of toepassing van (resultaten van) het voorspelmodel in beleidsdocumenten, zoals een beleidsplan? <ul style="list-style-type: none"> • Hoe?

Bijlage 6 Initiële codeboom

▼ ● ◇	1. Doorwerking	0
● ◇	Doorwerkign in beleidshandelingen	0
● ◇	Doorwerking in beleidsdocumenten	0
● ◇	Initiator	0
● ◇	Nageschakeld	0
● ◇	2. Input legitimiteit inhoudelijk	0
● ◇	3. Input legitimiteit procedureel	0
● ◇	4. Throughput legitimiteit inhoudelijk	0
● ◇	5. Throughput legitimiteit procedureel	0
● ◇	6. Output legitimiteit inhoudelijk	0
● ◇	7. Output legitimiteit procedureel	0
● ◇	8. Feedback legitimiteit inhoudelijk	0
● ◇	9. Feedback legitimiteit procedureel	0

Bijlage 7 Finale codeboom

▼	1. Doorwerking	48
	Betrokkenheid VM	6
	Doorwerking in beleidsdocumenten	12
	Doorwerking in beleidshandelingen	24
	Initiator	3
	Nageschakeld	4
▼	2. Input legitimiteit inhoudelijk	60
	Belang kwaliteit data	9
	Brede kenmerken	13
	Gemeente afhankelijk	4
	Overwegingen (waarden, visie)	21
	Verschillende perspectieven	6
	Voorspellende waarde	9
	Wens recentere data	2
▼	3. Input legitimiteit procedureel	36
	Bewustzijn	2
	Overwegingen	14
	Samenwerken regionaal	6
	Samenwerking verschillende functies	14
	Vertrouwen in proces	2
▼	4. Throughput legitimiteit inhoudelijk	21
	Kennis over algoritme	9
	Koppeling lokale data	3
	Niet complex	1
	Overwegingen (voorspelfout)	6
	Samenwerking data & beleid	2
▼	5. Throughput legitimiteit procedureel	15
	Overwegingen	9
	Vertrouwen in proces (betrokkenheid experts)	7

▼	6. Output legitimiteit inhoudelijk	48
	(te) algemene voorspelling	20
	Externe omstandigheden	8
	Niet leesbaar	2
	Onzekerheid	15
	Representatief	4
	Vertrouwen	2
▼	7. Output legitimiteit procedureel	127
	Advies / waarschuwing	16
	Belang bewustzijn	8
	Belang interpreteren	31
	Belang management	4
	Belang routine/structuur	4
	Belang verifiëren met praktijk	2
	Gebruik aanwakkeren	8
	Gebruik andere motieven	1
	Onderbouwen	16
	Overwegingen	10
	Rol individu	4
	Training	9
	Uitlegbaarheid	3
	Vergelijken met eigen cijfers	9
	Vergelijken wijken	1
	Vrijblijvend	6
	Weinig / beperkt gebruik	16

▼	8. Feedback legitimiteit inhoudelijk	18
	Actualisering model	6
	Feedback eigen cijfers	2
	In overleg met data experts	3
	Kritische vragen	3
	Onrealistische wensen	1
	Wensen	3
▼	9. Feedback legitimiteit procedureel	41
	Doorontwikkelen	8
	In ontwikkeling blijvend	14
	Overwegingen	4
	Samenwerking	14
	Verantwoording afleggen	3

▼	10. Mogelijkheden VM	76
	Aannemen personeel	1
	Beleidsvoorbereiding	1
	Beschrijving model	1
	Beter iets dan niets	9
	Capaciteitsberekeningen	6
	Check met eigen cijfers	4
	Datagedreven werken	14
	Efficiëntie	3
	Grip krijgen op toekomst	23
	Ijkpunt	2
	Inzet middelen	4
	Inzicht in kosten	9
	Lokaal inzoomen	5
	Meer bewustwording (doorontwikkeling)	1
	Politiek onderbouwen	2
	Preventie	1
	Scenario's berekenen	2
	Strategische planning uitvoering	4
▼	11. Beperkingen VM	21
	Knoppen draaien beperkt	4
	Lokaal inzoomen beperkt	15
	Niet aan de praat krijgen	2
	Vergelijking tussen gemeenten	1
	Wendbaarheid	1
▼	12. Risico's VM	25
	Betrouwbaarheid	22
	Goede data	4
	13. Belang capaciteit	8
	14. Belang expertise / kennis	34
	15. Humans in the loop	7
	Rotterdamse model	5
	Voorgeschiedenis	25