



Onderzoek

Dataverzameling en -deling voor onderzoek naar online kansspelen

Auteurs

Ir. Tommy van der Vorst

Melvin Hanswijk MSc LLM

Prof. dr. mr. Gerrit-Jan Zwenne

Ir. Laura Smulders

Onderzoek

Dataverzameling en -deling voor onderzoek naar online kansspelen

Auteurs

Ir. Tommy van der Vorst
Melvin Hanswijk MSc LLM
Prof. dr. mr. Gerrit-Jan Zwenne
Ir. Laura Smulders
m.m.v. Joost Crielaard BSc

Opdrachtgever

Wetenschappelijk Onderzoek- en Datacentrum (WODC)

Begeleidingscommissie

Prof. mr. dr. C.J. Wolswinkel (voorzitter)
Dr. J.H. Hoepman
Mr. dr. M.H. Paapst
Drs. B.J.M. Leushuis (lid; beleidsdirectie)
Dr. P. van de Ven (lid; projectbegeleider WODC)

Publicatienummer

2024.164-2525

Citeren als

Dialogic: van der Vorst, T., Hanswijk, M., Zwenne, G., Smulders, L. (2025). *Dataverzameling en -deling in verband met onderzoek naar online kansspelen*. In opdracht van het WODC.

Datum

9 oktober 2025

Beeld omslag

Anna Shvets, Pexels License

© 2025, Dialogic innovatie & interactie. Auteursrechten voorbehouden. Niets uit dit rapport mag worden verveelvoudigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm, digitale verwerking of anderszins, zonder voorafgaande schriftelijke toestemming van Dialogic innovatie & interactie.

Inhoud

Managementsamenvatting	4
1 Inleiding	7
1.1 Doelstelling en onderzoeksvragen	7
1.2 Onderzoeksaanpak	11
2 Onderzoeks- en databehoefte	15
2.1 Soorten onderzoek	15
2.2 Soorten onderzoekers	17
2.3 Huidige invulling en knelpunten	20
3 Juridische analyse	25
3.1 Juridische probleemstelling	25
3.2 Borgen van rechtmatigheid datadeling	28
3.3 Borgen van medewerking	45
3.4 Taakstelling Ksa en doelstelling CDB	54
3.5 Conclusie	55
4 Oplossingsrichtingen	60
4.1 Analyse kader	60
4.2 Oplossingsrichtingen	62
4.3 Afweging	80
5 Conclusie	84
5.1 Beantwoording onderzoeksvragen	84
Verwijzingen	91
Bijlage 1. Overzicht interviewrespondenten en gesproken personen	94
Bijlage 2. Casuïstiek	96
Inventarisatie van casussen	96
Selectie van casussen	96
Analyse van geselecteerde casussen	99

Managementsamenvatting

Het doel van dit onderzoek is om de behoefte aan spelersdata voor onafhankelijk onderzoek met maatschappelijk en/of wetenschappelijk doel op het gebied van speelgedrag bij kansspelen te inventariseren. We onderzoeken of momenteel in deze behoefte wordt voorzien en zo niet, wat daarvoor de belangrijkste technische, juridische en organisatorische knelpunten zijn. Op basis hiervan worden oplossingsrichtingen gevormd, waarmee op termijn in (een groter deel van) de behoefte zou kunnen worden voorzien.

Onderzoeks- en databehoeft

Op dit moment zijn vergunde kansspelaanbieders wettelijk verplicht om geanonimiseerde data te delen met onderzoekers die daar bilateraal om vragen ten behoeve van onderzoek naar kansspelverslaving (Wet op de kansspelen (Wok) art. 31m). Deze voorziening is nog beperkt gebruikt en kent een aantal knelpunten: (1) het proces is suboptimaal ingericht, (2) vanwege de vereiste anonimisering kunnen de gegevens niet op spelersniveau worden gekoppeld tussen aanbieders, (3) vanwege de anonimisering kunnen geen achtergrondkenmerken worden gekoppeld aan de gegevens, (4) de Kansspelautoriteit (Ksa) kan zelf niet alle onderzoeken uitvoeren die het zou willen, en (5) de toegang op basis van het wetsartikel is beperkt tot onderzoek ter preventie van kansspelverslaving. De Ksa verzamelt weliswaar gestructureerde data ten behoeve van toezicht in een controledatabank, maar kan deze data niet zonder meer doorgeven aan onderzoekers, noch zelf voor alle onderzoeken gebruiken.

Juridische obstakels

Er zijn op drie gebieden juridische obstakels die verdergaande datadeling mogelijk in de weg staan:

1. **Rechtmatigheid van het delen van data.** De AVG verbiedt het verwerken van bijzondere persoonsgegevens (waarvan hier naar verwachting sprake is). Om hiermee om te gaan moeten gegevens worden geanonimiseerd (evt. na koppeling) of moet gebruik worden gemaakt van een uitzondering voor wetenschappelijk onderzoek (UAVG art. 24). Daarnaast kan de kansspelwetgeving beperkend werken ten aanzien van het delen van niet-geanonimiseerde gegevens voor andere doelen dan kansspelverslavingspreventie.
2. **Borgen van de medewerking van aanbieders.** Er kan niet zonder meer worden uitgegaan van het vrijwillig delen van data door aanbieders, buiten de geldende verplichting. Mogelijke oplossingen zijn (1) zelfregulering, (2) het aanpassen van de kansspelwetgeving, of (3) het CBS per AMvB de opdracht geven de data te verzamelen.
3. **Taakstelling van de Ksa.** De Ksa kan alleen onderzoek uitvoeren dat aansluit bij haar wettelijke taken. Gegeven de huidige taakstelling kan de Ksa echter niet

alle onderzoeken uitvoeren die zij zou willen uitvoeren. Om dit toch mogelijk te maken, moet de taakstelling van de Ksa in artikel 33b Wok worden aangepast.

Oplossingsrichtingen

Er zijn vier oplossingsrichtingen geïdentificeerd voor verdergaande datadeling, waarmee de bovengenoemde knelpunten kunnen worden opgelost (rekening houdend met de juridische kaders):

1. **De huidige bilaterale werkwijze kan worden verbeterd.** Door de rol van de Kansspelautoriteit in het proces te vergroten, kan het proces soepeler verlopen. Er wordt echter niet voorzien in betere koppelbaarheid van spelersdata.
2. **Er kan een decentraal platform worden ingericht voor datadeling.** Toegang tot de data en waarborgen ten aanzien van herleidbaarheid worden grotendeels technisch geborgd. Deze oplossing biedt een hoge mate van veiligheid, maar is relatief complex en vraagt investeringen van de aanbieders.
3. **Er kan een centrale intermediair worden aangesteld,** die datadeling faciliteert en koppelbare pseudoniemen voor spelers verzorgt. Dit zou een derde partij of de Kansspelautoriteit (na wijziging taakstelling) kunnen zijn. Deze oplossing leunt sterker op organisatorische maatregelen.
4. **Het CBS kan de data verzamelen en beschikbaar stellen via CBS Microdata.** Deze route maakt gebruik van de CBS-bevoegdheid voor het verzamelen van de data, en de bestaande faciliteiten en kaders voor werken met microdata.

Aansluiting bij het datamodel dat wordt gehanteerd voor toezicht door de Ksa ligt in alle gevallen voor de hand (CDB/GDT). De oplossingsrichtingen verschillen ten aanzien van de juridische route, benodigde investeringen, realisatietermijn, en risico's ten aanzien van gegevensbescherming. De afweging tussen kosten en baten is niet in dit onderzoek gemaakt, maar uiteraard relevant.

Bij ontwikkeling van beleid kan ervoor worden gekozen om een oplossingsrichting ten dele, of gefaseerd in tijd, te effectueren. Het is bijvoorbeeld denkbaar dat op korte termijn de deeloplossingen ter verbetering van het proces worden ingezet, terwijl de koppelbaarheid met achtergrondgegevens pas op langere termijn wordt gerealiseerd. De keuze voor een bepaalde oplossingsrichting kan echter wel bepalend zijn voor hoe verbreding van de datadeling in de toekomst kan gebeuren.

Onderzoeksmethode

Het onderzoek bestond uit drie fasen: (1) verkenning, (2) verdieping en (3) validatie. In de verkennende fase zijn voor de behoefteverkenning interviews uitgevoerd met onderzoekers en gebruik gemaakt van literatuur. De juridische verkenning is gebaseerd op de relevante wetsteksten, jurisprudentie en een interview. In de verdiepende fase is met interviews en bureauonderzoek nader gekeken naar een aantal bestaande vergelijkbare casussen, die zijn gevonden en geselecteerd in de verkenning volgens een

selectiekader. Er hebben daarnaast gesprekken plaatsgevonden met de Ksa, het CBS en de Autoriteit Persoonsgegevens. In de validatiefase is een sessie georganiseerd met belanghebbenden, waarbij een kansspelaanbieder, onderzoeker, de Ksa, het CBS en het ministerie van J&V vertegenwoordigd waren. Een begeleidingscommissie zag tot slot toe op de kwaliteit van het onderzoek en de eindproducten.

1 Inleiding

Op verzoek van het Wetenschappelijk Onderzoek- en Datacentrum (WODC) van het ministerie van Justitie en Veiligheid heeft Dialogic innovatie & interactie (hierna: Dialogic) in samenwerking met Gerrit-Jan Zwenne onderzoek uitgevoerd rondom het verzamelen en beschikbaar stellen van data over spelersgedrag bij online kansspelen ten behoeve van onderzoek.

Leeswijzer

In dit eerste hoofdstuk gaan we in op de aanleiding van het onderzoek, de doelstellingen en onderzoeksvragen, en de onderzoeksaanpak. In hoofdstuk 2 beschrijven we de behoefte aan data vanuit onderzoek, welk deel daarvan niet wordt ingevuld, en wat de voornaamste knelpunten zijn op technisch en organisatorisch gebied daarbij. In hoofdstuk 3 gaan we in op het juridisch kader, en gaan we in op juridische knelpunten bij bredere deling van data, en oplossingen daarvoor. In hoofdstuk 4 vertalen we de bevindingen naar oplossingsrichtingen, waarmee de beschikbaarheid van data ten behoeve van onderzoek kan worden verbeterd. Hoofdstuk 5 sluit af met een conclusie, waarin de onderzoeksvragen worden beantwoord.

In Bijlage 1 is een overzicht te vinden van respondenten. In Bijlage 2 gaan we nader in op de casussen die we in het kader van dit onderzoek hebben bekeken.

In de tekst zijn verwijzingen naar literatuur gegeven als cijfers tussen [blokhaken]. Op pagina 91 zijn deze bronnen te vinden. Cijfers in superscript verwijzen naar voetnoten onderaan de pagina.

1.1 Doelstelling en onderzoeksvragen

Sinds de legalisering van de online kansspelmarkt in 2021 zijn er signalen van zorgwekkende ontwikkelingen op het gebied van onvoldoende consumentenbescherming en verslavingspreventie. [Min233] Om maatregelen ter verbetering daarvan te treffen is een goede probleemanalyse nodig, en moeten signalen in de brede context van het kansspelbeleid kunnen worden geduid. Mede daarom is er behoefte om het gedrag van spelers en de effecten van kansspelen daarop te kunnen monitoren. En om dat goed te kunnen doen, zijn gegevens op spelersniveau nodig.

Het doel van dit onderzoek is om de behoefte aan spelersdata ten behoeve van onafhankelijk onderzoek met maatschappelijk en/of wetenschappelijk doel op het gebied van speelgedrag bij kansspelen te inventariseren. Op basis hiervan onderzoeken we of in deze behoefte wordt voorzien en zo niet, wat daarvoor de belangrijkste technische, juridische en organisatorische knelpunten zijn. Op basis hiervan worden

oplossingsrichtingen gevormd, waarmee op termijn in (een groter deel van) de behoefte zou kunnen worden voorzien.

De volgende onderzoeksvragen staan centraal in het onderzoek:

1. Hoe ziet op hoofdlijnen de behoefte aan onafhankelijk onderzoek met maatschappelijk en/of wetenschappelijk doel op het gebied van speelgedrag bij kansspelen op afstand eruit, en wat is de daaruit voortvloeiende databehoeftes ten aanzien van beschikbare spelersdata?
2. Wat zijn op dit moment de juridische, technische en praktisch-organisatorische mogelijkheden tot het verstrekken van spelersdata om in de behoefte te voorzien?
3. Wat zijn op dit moment de juridische, technische en praktisch-organisatorische obstakels met betrekking tot het verstrekken van spelersdata om verder in de behoefte te voorzien?
4. Welke oplossingsrichtingen zijn er voor het beschikbaar stellen van spelersdata om verder (dan nu mogelijk is) in de onderzoeksbehoefte te kunnen voorzien?

1.1.1 Afbakening

Hieronder gaan we nader in op de interpretatie van enkele begrippen die centraal staan in het onderzoek en de onderzoeksvragen: (1) aanbieders, (2) onafhankelijk onderzoek, (3) maatschappelijk doel, (4) databehoeftes en (5) beschikbare data.

Aanbieders

Waar wordt gesproken over (kansspel)aanbieders worden in dit rapport de partijen bedoeld met een vergunning om in Nederland online kansspelen aan te bieden.^a

Onafhankelijk onderzoek

Met onafhankelijk onderzoek bedoelen we onderzoek waarvan noch het al dan niet plaatsvinden, noch de uitkomsten of conclusies kunnen worden beïnvloed door belanghebbenden.

Bij onafhankelijk onderzoek staat waarheidsvinding voorop. Het wegen van belangen van belanghebbenden is een afweging die vervolgens, op basis van de onderzoeksresultaten, op politieke basis en in publiek debat kan worden gemaakt. Onafhankelijkheid is strikt genomen een eigenschap van de wijze waarop een onderzoek wordt uitgevoerd. Het is echter ook van groot belang dat de *onderzoeker* en de *onderzoeks-*

^a Omdat sprake is van een open vergunningenstelsel staat het aantal partijen dat een vergunning exploiteert niet vast. Op 31 januari 2025 waren er dertig vergunningen voor online kansspelen afgegeven, waarbij een bedrijf meerdere vergunningen kan hebben. [4]

instelling waarmee deze verbonden is onafhankelijk zijn, omdat anders (onbewust) *bias* kan optreden, en ook sturing in het onderzoeksproces niet is uit te sluiten.

Naast invloed op de uitkomsten van een onderzoek is ook invloed op de vraag of onderzoek al dan niet wordt uitgevoerd, en al dan niet wordt gepubliceerd, van groot belang. Hoewel het denkbaar is dat belanghebbenden onafhankelijk onderzoek laten uitvoeren, is onwenselijk dat alleen onderzoek wordt uitgevoerd of gepubliceerd waarvan de resultaten alleen of grotendeels in het belang van deze belanghebbenden zijn.

Maatschappelijk doel

Met 'maatschappelijk doel' wordt verwezen naar onderzoeken die worden uitgevoerd ten behoeve van de maatschappij. Onderzoek dat met publieke middelen wordt gefinancierd of door publieke organisaties wordt uitgevoerd, is dat haast per definitie – er is in die gevallen (uiteindelijk politiek) besloten dat een onderzoek maatschappelijke meerwaarde biedt. Hierbij gaat het primair om wetenschappelijk onderzoek (waarbij het doel is om kennis te vergaren) en beleidsonderzoek (waarbij het doel is om kennis te vergaren om uiteindelijk beleid te kunnen vormgeven).

Databehoefte

Met *databehoefte* verwijzen we naar de gegevens die nodig zijn om onderzoek van het hierboven voorziene soort uit te kunnen voeren. In de praktijk zijn voor het beantwoorden van een onderzoeksvraag vaak verschillende onderzoeksmethoden in te zetten en te combineren. De meest relevante databehoefte in dit onderzoek is dan ook de data waarvoor op dit moment geen (goed) alternatief bestaat.

De databehoefte betreft niet alleen de vraag aan *welke* gegevens behoefte is, maar ook over de wijze *waarop* deze beschikbaar worden gesteld en de processen daaromheen. Daarnaast kan de tijdsdimensie van belang zijn: bestaat er een bepaalde eenmalige databehoefte, of gaat het bijvoorbeeld om data ten behoeve van langer lopend longitudinaal onderzoek of langlopende monitoring? Hoe vaak dienen de gegevens verversd te worden, en hoe actueel moeten de gegevens zijn?

Beschikbare spelersdata

Met beschikbare spelersdata bedoelen we een specifieke set gegevens waarover aanbieders van online kansspelen op spelersniveau momenteel moeten beschikken op grond van de geldende wet- en regelgeving:

- Gegevens die door aanbieders moeten worden verzameld ter signalering van onmatig speelgedrag (art. 31m Wet op de Kansspelen (Wok) en art. 13-14 Besluit Werving, reclame en verslavingspreventie kansspelen (Wrvk));
- Gegevens die moeten worden vastgelegd in de Controledatabank (CDB) (art. 5.3 Besluit Kansspelen op Afstand (Koa));
- Gegevens die worden verzameld ter identificatie van de speler (art. 4.11 Besluit Koa).

Het is uiteraard mogelijk dat er een onderzoeksbehoefte is waarin met alleen deze beschikbare spelersdata *niet* kan worden voorzien, omdat deze data daarvoor niet toereikend is. Als de beschikbare data slechts minimaal in de bestaande behoefte kan voorzien is dat een relevante constatering, omdat dan mogelijk andere data verzameld moeten worden. De focus van het huidige onderzoek is echter om de behoefte in kaart te brengen waarin met bij de aanbieders beschikbare spelersdata kan worden voorzien. Het is relevant om te overwegen in hoeverre een oplossingsrichting voor het beschikbaar stellen van deze data voor onderzoeksdoeleinden in de toekomst zou kunnen worden uitgebreid met aanvullende data.

Hieronder volgt een overzicht van de specificaties van de bovengenoemde categorieën beschikbare spelersdata.

Gegevens over onmatig speelgedrag

Een vergunninghouder is gehouden om van iedere speler gegevens met betrekking tot het speelgedrag te registreren en te analyseren, om onmatig speelgedrag^a tijdig te signaleren, op grond van art. 31m Wok en art. 13-14 Besluit Wrvk.^b Er moeten in ieder geval gegevens worden verzameld met betrekking tot inschrijving, aanmelding, leeftijd, duur van deelname, frequentie van deelname, grenzen van het speelgedrag, aanpassingen en overschrijdingen van deze grenzen, inzetten en speluitkomsten, betalingstransacties, mutaties op de speelrekening en schorsing van de mogelijkheid tot deelname.^c

Gegevens ten behoeve van controle (CDB)

De Ksa heeft ten behoeve van haar toezichthoudende taak toegang tot gegevens van de vergunde aanbieders. De aanbieders zijn verplicht om specifieke gegevens bij te houden in een controledatabank (CDB) en deze doorlopend op te leveren aan de Ksa. De Ksa aggregereert deze gegevens in de *Geaggregeerde Databank Toezicht* (GDT). Los hiervan beschikt de Ksa over de mogelijkheid om gegevens op te vragen bij de aanbieders.

Welke gegevens in de CDB moeten worden vastgelegd, wordt bepaald in artikel 5.3 van het Besluit Koa:^d

- een niet direct tot de identiteit van de speler herleidbare, unieke aanduiding van iedere speler;

^a Onder *onmatig speelgedrag* wordt risicovol speelgedrag verstaan dat kan leiden tot kansspelsverslaving (Art. 2 lid 1 Besluit Wrvk) [6]

^b Art. 13 Besluit Wrvk [6]

^c Art. 14 Besluit Wrvk [6]

^d Art. 5.3 Besluit Koa [35]

- de datum, het tijdstip en de aard van iedere aanpassing en overschrijding van het spelersprofiel;
- de aanleiding voor en de datum, het tijdstip en de aard van iedere interventie als bedoeld in artikel 18 van het Besluit werving, reclame en verslavingspreventie kansspelen;
- de datum, het tijdstip, het bedrag en de aard van iedere creditering en iedere debitering van iedere speelrekening, alsmede, waar van toepassing, de aard van het daarbij gebruikte betaalinstrument en, waar van toepassing, een niet direct tot de identiteit van de speler herleidbare, unieke aanduiding van de daarbij gedebiteerde onderscheidenlijk gecrediteerde betaalrekening,
- de totale inzet en speelwinst per spelsoort en, voor zover het betreft een kansspel waarbij de vergunninghouder anders dan als inzet een vergoeding ontvangt voor het geven van gelegenheid tot deelneming aan dat kansspel, het totale bedrag van die vergoeding.

Gegevens ter identificatie van spelers

Een vergunninghouder mag een persoon alleen inschrijven als speler wanneer deze beschikt over bepaalde persoonsgegevens ter identificatie van de speler op grond van art. 4.11 Besluit Koa. In het Besluit Koa is vastgelegd om welke gegevens dit minimaal gaat, en in welke mate deze gevalideerd moeten worden.^a De vergunninghouder schrijft een persoon slechts in als speler, indien hij beschikt over ten minste:^b

- diens naam, voornamen, geboortedatum, geboorteplaats en, voor zover die persoon daarover beschikt, diens burgerservicenummer (BSN), zoals die zijn opgenomen op het identiteitsdocument van die persoon;
- diens fysieke adres en contactgegevens, waaronder in ieder geval e-mail adres en telefoonnummer, en
- het rekeningnummer of een andere unieke aanduiding en de tenaamstelling van de tegenrekening.

Het BSN dient na gebruik voor de eerste identificatie en registratie te worden verwijderd door de aanbieder.^c

1.2 Onderzoeksaanpak

Het onderzoek bestaat uit drie fasen: (1) verkenning, (2) verdieping en (3) validatie.

^a Art. 4.11 Besluit Koa [16]

^b Art. 4.11 Besluit Koa [16]

^c Art. 20 Besluit Wrvk [4]

Verkenning

In de eerste fase van het onderzoek worden twee aspecten van de onderzoeksvraag verkend:

1. De behoefte aan data over speelgedrag ten behoeve van onderzoek die op dit moment niet wordt ingevuld;
2. De knelpunten op technisch en organisatorisch vlak die daarbij een rol spelen;
3. Mogelijke oplossingsrichtingen op technisch en organisatorisch vlak (verkenning van vergelijkbare casussen van datadeling);
4. Juridische knelpunten en oplossingen daarvoor (juridisch kader).

In deze fase is zowel gebruik gemaakt van literatuur en bestaande databronnen als nieuw verzamelde data, voornamelijk via interviews. Hieronder lichten we de gebruikte onderzoeksmethoden toe.

Literatuurstudie is ingezet voor de afbakening van de centrale begrippen in dit onderzoek (zoals de specificatie van beschikbare spelersdata), het opstellen van het juridisch kader over het (her)gebruik van data voor onderzoek, en het selecteren van vergelijkbare casussen waarin een overheidsinstantie gegevens verzamelt op persoonsniveau en deze beschikbaar stelt voor onafhankelijk onderzoek. Het heeft inzicht gegeven in hoe vergelijkbare kwesties over het delen en gebruiken van data in andere sectoren zijn opgelost.

Er zijn verkennende interviews uitgevoerd met onderzoekers en organisaties die onderzoek doen naar (online) kansspelverslaving. Het doel was om zo breed mogelijk ('divergerend') de (mogelijke) onderzoeksbehoefte in kaart te brengen. De besproken thema's omvatten de onderzoeks- en databehoefte, de noodzaak om spelers te koppelen, het koppelen van achtergrondkenmerken, het tijdsperspectief van de onderzoeken en governance. Van de zes interviews hebben we vier verschillende onderzoekers/organisaties daadwerkelijk gesproken, terwijl we van de overige twee onderzoekers/organisaties schriftelijk antwoorden op onze vragen hebben ontvangen. Voor het opstellen van het juridisch kader is er één interview met een overheidsinstantie afgenomen. Hierbij hebben we informatie opgehaald over de juridische mogelijkheden om zowel de medewerking als de rechtmatigheid van de gegevensverwerking te waarborgen. De (gemiddelde) duur van de verkennende interviews bedroeg een uur.

Voor de casuïstiek is via literatuurstudie en op basis van de interviews een groslijst gemaakt van mogelijk relevante casussen. Op basis van een selectiekader is vervolgens een aantal casussen geselecteerd voor nadere analyse. De analyse vond plaats op basis van een analysekader. Zie voor nadere toelichting op deze casestudies en het selectie- en analysekader de uitwerking in Bijlage 2.

Verdieping

In de verdiepende fase van het onderzoek zijn op basis van de verkenning oplossingsrichtingen geformuleerd. Deze zijn nader aangescherpt op basis van verdiepende gesprekken en nadere analyse. De verdiepende gesprekken zijn uitgevoerd met vertegenwoordigers van kansspelaanbieders en betrokkenen bij vergelijkbare datadelingen in casestudies. Daarnaast is gesproken met de Kansspelautoriteit (Ksa) en de Autoriteit Persoonsgegevens (AP).^a In acht interviews hebben we veertien personen gesproken. De verdiepende gesprekken duurden gemiddeld een uur.

De besproken thema's omvatten onder andere de benodigde tijd om de oplossingsrichting te realiseren, de kosten, het draagvlak en de bereidheid tot medewerking, vereiste wetswijzigingen, ethische bezwaren, technische en praktische knelpunten, de borging van betrouwbaarheid en volledigheid van aangeleverde data, evenals de toegang en herleidbaarheid door aanbieders.

Validatie

Tot slot is een validatiesessie uitgevoerd op basis van opgedane kennis. In deze sessie zijn eerdere bevindingen uit de juridische analyse en interviews gedeeld om vervolgens feedback op te halen van belanghebbenden ten aanzien van de oplossingsrichtingen. De besproken thema's omvatten onder meer de voorlopige resultaten van de onderzoeksbehoefte, de juridische analyse en de oplossingsrichtingen. Respondenten kregen de ruimte om feedback te geven over de doelmatigheid en doeltreffendheid, evenals op de voor- en nadelen van de mogelijke oplossingsrichtingen. De duur van de validatiesessie was twee uur.

Bijlage 1 geeft een overzicht van alle gesprekspartners (in de verkennende en verdiepende fase) en deelnemers aan de validatiesessie.

Triangulatie en weging

De boven beschreven onderzoeksaanpak maakt gebruik van verschillende onderzoeksmethoden die worden gecombineerd ('triangulatie') om tot betrouwbare en onderbouwde bevindingen te komen. In het onderzoek worden bronnen en (mondelinge en schriftelijke) bijdragen van belanghebbenden als bron gebruikt. Deze bijdragen en bronnen zijn zoveel mogelijk verzameld op een vooraf vastgestelde gestructureerde wijze (zo werd voor de diverse gesprekken steeds een interviewprotocol en/of 'topiclijst' gehanteerd). De bijdragen zijn door de onderzoekers gewogen in het licht van het belang van de belanghebbende. Er is daarnaast rekening gehouden met het feit dat respondenten (om verschillende redenen) in sommige gevallen niet beschikbaar waren

^a De Autoriteit Persoonsgegevens is toezichthouder op grond van privacywetgeving. In het onderzoek is op informele basis en aan de hand van een concept van de juridische analyse van gedachten gewisseld met de AP. De deelname van AP dient niet te worden geïnterpreteerd als goedkeuring of instemming met de in het onderzoek weergegeven juridische standpunten.

voor deelname. Tussentijdse bevindingen zijn in een validatiesessie voorgelegd aan een vertegenwoordiging van een selectie van belanghebbenden en onafhankelijke experts (waarvan de meesten ook hebben deelgenomen aan de verkennende fase). Diverse belanghebbenden, waaronder meerdere kansspelaanbieders zijn in de gelegenheid gesteld deel te nemen aan interviewgesprekken en de validatiesessie.

Onafhankelijkheid en kwaliteitsborging

De onderzoekers zijn in staat gesteld om het onderzoek onafhankelijk uit te voeren en hierover te rapporteren. Dialogic, de betrokken onderzoekers, noch de leden van de begeleidingscommissie hebben enig belang bij de uitkomsten van dit onderzoek.

Een door het WODC ingestelde begeleidingscommissie heeft gedurende het onderzoek toegezien op de onafhankelijkheid en kwaliteit ervan. Onder andere de onderzoeksmethoden, interviewprotocollen en voorstellen voor respondenten(selectie) zijn aan de begeleidingscommissie voorgelegd en zo nodig bijgesteld. Daarnaast heeft de begeleidingscommissie conceptversies van dit rapport en tussenresultaten ontvangen en een aantal keer de mogelijkheid gekregen om hier inhoudelijk op te reageren. De onderzoekers benadrukken geen sturing te hebben ervaren vanuit de begeleidingscommissie danwel de opdrachtgever ten aanzien van de onderzoeksresultaten. De onderzoekers danken de leden van de begeleidingscommissie, alle interviewrespondenten en deelnemers aan de validatiesessie voor hun waardevolle bijdragen aan dit onderzoek.

2 Onderzoeks- en databehoeftes

In dit hoofdstuk brengen we de behoefte aan data op spelerniveau ten behoeve van onderzoek over speelgedrag bij kansspelen in kaart. In het bijzonder gaat het daarbij om de behoefte die op dit moment niet of onvoldoende wordt ingevuld, en de organisatorische en technische knelpunten die daaraan ten grondslag liggen.

De behoefte aan data kan voortvloeien uit onderzoeksbehoefte, maar beschikbaarheid van data kan ook leiden tot activering van latente onderzoeksbehoefte. In dit onderzoek starten we primair vanuit de *onderzoeksbehoefte*. In §2.1 gaan we in op de verschillende *soorten onderzoek* en de daaruit voortvloeiende behoefte aan data. In §2.2 gaan we in op verschillende soorten *onderzoekers* die dergelijk onderzoek uitvoeren. In §2.3 bespreken we knelpunten op organisatorisch en technisch vlak, en wensen ten aanzien van (tot nu toe onvervulde) onderzoeks- en databehoeftes.

2.1 Soorten onderzoek

Op basis van de verkenning onderscheiden we twee (ten aanzien van de databehoeftes relevant te onderscheiden) onderzoekscategorieën. Deze tweedeling helpt om beter inzicht te krijgen in de verschillende doelstellingen, methoden en databehoeften die voor kansspelonderzoek relevant zijn, en lijkt (op basis van de gesprekken) bepalend te kunnen zijn voor de uiteindelijke oplossingsrichting.

Gezondheidswetenschappelijke of public health-onderzoeken

Gezondheidswetenschappelijke of *public health*-onderzoeken richten zich op het in kaart brengen van gokgerelateerde problematiek op macroniveau. Deze onderzoeken monitoren en analyseren trends, risicogroepen en de prevalentie van probleemgedrag onder spelers. Dit is een fundamentele benadering om het ontstaan van speelgedrag en probleemgokken te begrijpen met of zonder de aanwezigheid van ander risicovol gedrag (bijv. alcoholgebruik, delinquentie).

De focus binnen dit type onderzoek ligt op geaggregeerde statistieken, waarbij het speelgedrag van grote groepen spelers wordt samengevat en geanalyseerd. Denk hierbij aan het totale aantal probleemspelers in een bepaalde periode, of het verloop van risicogedrag over tijd. Hoewel de benadering op macroniveau ligt, is er wel behoefte aan gedetailleerde uitsplitsingen. Zo is er de behoefte om ook de rol van voorspellers in meerdere domeinen te onderzoeken, waaronder neurocognitieve, socio-demografische, psychologische en gedragsfactoren, over een lange periode. Er is om die reden data nodig op spelersniveau.

Public health-onderzoeken kennen zowel longitudinale onderzoeken, waarbij spelersgedrag over langere tijd wordt gevolgd, als cross-sectionele onderzoeken, gericht op specifieke risicofactoren of onderzoek naar gedragsindicatoren zoals het zogeheten

markers of risk-onderzoek. [Kan251] Hierbij geeft de Ksa aan dat problematisch speelgedrag niet uitsluitend gekenmerkt wordt door extreme verliezen, maar dat een bredere set indicatoren nodig is, zoals speelfrequentie, aantal stortingen per dag, en nachtelijk speelgedrag. Een voorbeeld van public health onderzoeken is het maken van een profielschets van probleemspelers en hun speelgedrag. Op basis van CDB-data en aanvullende achtergronddata zouden risicoprofielen kunnen worden opgesteld.

Relevant is om op te merken dat onderzoek in deze categorie in mindere mate *‘in de spellen zelf kijkt’*. De variabelen op het laagste niveau betreffen hier ‘inputs’ en ‘outputs’ van een kansspel, maar bijvoorbeeld geen informatie over onderliggende gedragingen (zoals chatberichten, bewegingen van de muis, et cetera). In een dataset kunnen gegevens van verschillende spellen en varianten daarvan voorkomen en worden samengevoegd. Het betreft hier, anders gezegd, de *generieke* kenmerken die iedere aanbieder van online kansspelen zou moeten kunnen opleveren, ongeacht de specifieke spelvorm en invulling.

Gedrags- en neurologische onderzoeken

Gedrags- en neurologische onderzoeken op microniveau richten zich op het analyseren van de impact van specifieke spelkenmerken of spelersgedragingen op kansspelverslaving. Dit is een toegepaste benadering die zich richt op het beter begrijpen van verantwoord speelgedrag en de rol die specifieke spelontwerpen daarin spelen. Gedrags- en neurologische onderzoeken hanteren een ontwerpbenadering voor verantwoord gokken (of beleid) door een kader van gepersonaliseerde RG (Responsible Gaming) tools te ontwerpen (bijv. limieten, educatieve interventies, geplande pauzes, verschillende soorten en vormen van communicatie met spelers), om hun doeltreffendheid te vergroten door dergelijke maatregelen af te stemmen op de behoeften en voorkeuren van spelers.

Voor de interpretatie van de data voor dit soort onderzoek is actieve medewerking van kansspelaanbieders essentieel. De onderzoekers hebben toegang nodig tot gebruikersdata en spelstatistieken. Daarnaast fungeert het platform als onderzoeksomgeving, waarbij experimentele aanpassingen aan het spel mogelijk vereist zijn. Onderzoeken zijn incidenteel van aard. In sommige gevallen zijn onderzoeken (deels) invasief, omdat ze actieve interventies of metingen vereisen. Een voorbeeld van gedrags- en neurologisch onderzoek op microniveau is het testen van hypothesen over de impact van sociale interactie via chatfuncties binnen een spel op verslavingsgedrag. De resultaten van dit soort onderzoeken kunnen bijdragen aan het inzichtelijk maken van risicofactoren of verbeterde maatregelen binnen het spelontwerp.

Overzicht

Onderstaande Tabel 1 toont de belangrijkste eigenschappen en verschillen tussen de twee onderscheiden typen onderzoeken.

Tabel 1 Belangrijkste eigenschappen en verschillen tussen de twee onderscheiden typen onderzoeken met relevantie voor het voorliggende vraagstuk

	Gezondheidswetenschappelijke of public health-onderzoeken	Gedrags- en neurologische onderzoeken
Aard van de gegevens	Generiek, zelfde kenmerken voor alle aanbieders/spelen	Specifieke kenmerken per aanbieder en/of spel
Tijdsduur onderzoek	Veelal longitudinaal	Veelal eenmalig
Wijze van dataverzameling	Niet-invasief (alleen data die al beschikbaar is of kan worden verzameld zonder aanpassing aan het gokplatform/spel)	Soms ook invasief (een experiment uitvoeren binnen een online gokplatform)
Relevantie koppelbaarheid van spelersgegevens tussen platforms	Hoog, om te kunnen corrigeren voor overstapedrag en om een totaalbeeld te krijgen van een speler	Laag, wanneer onderzoek op één platform wordt uitgevoerd.
Belangrijkste knelpunten bij de huidige toegang tot data (via art. 31m lid 5 Wok)	Data is niet koppelbaar tussen aanbieders, medewerking en duiding aanbieders nodig. Onafhankelijkheid (c.q. perceptie daarvan) in het geding bij afspraken met aanbieders.	Medewerking aanbieders en specifieke afspraken zijn nodig bij meer diepgaande onderzoeken. Onafhankelijkheid (c.q. perceptie daarvan) in het geding bij afspraken met aanbieders.

2.2 Soorten onderzoekers

In het kader van dit onderzoek is gesproken met een aantal onderzoekers op het gebied van kansspelverslaving, verbonden aan academische of maatschappelijke instellingen. Deze onderzoekers geven (niet geheel onverwacht) aan dat toegang tot data voor wetenschappelijk en maatschappelijk onderzoek cruciaal is. In de gesprekken is een aantal aandachtspunten herhaaldelijk genoemd, waarbij met name (werkelijke of gepercipieerde) onafhankelijkheid van onderzoekers die werken met data van kansspelaanbieders wordt gezien als kritiek aspect.

Onderzoeken door of voor de kansspelsector

Allereerst kan de vraag worden gesteld in hoeverre kansspelaanbieders zélf onderzoek zouden moeten kunnen uitvoeren (met gebruikmaking van de data die zal worden gedeeld). Uit de gesprekken ontstaat het beeld dat dit in principe niet bezwaarlijk is – het

is immers in deze gevallen glashelder dat de aanbieder opdrachtgever is en de onderzoeksresultaten (danwel het überhaupt uitvoeren van het onderzoek) een bepaald belang kan dienen.

Hierbij speelt uiteraard de vraag in hoeverre dergelijk onderzoek in het belang is van de maatschappij of de wetenschap – wanneer de data bijvoorbeeld zou worden gebruikt ter optimalisatie van een spel of marketing van een kansspelaanbieder is er wellicht ook wetenschappelijk zinnige kennis ontstaan, maar zijn de maatschappelijke baten discutabel. Vanuit concurrentieel oogpunt kan het eveneens onwenselijk zijn wanneer kansspelaanbieders (via de voorgenomen datadeling) meer inzicht krijgen in hoe het er bij concurrenten in de bedrijfstak aan toegaat.

In het verlengde ligt de vraag in hoeverre onderzoekers *in opdracht van* kansspelaanbieders onderzoek zouden moeten kunnen uitvoeren op basis van de hier gedeelde data. In principe kan dit onderzoek betreffen dat relevant is voor beleidsvorming of de wetenschap. Het is ook niet ondenkbaar dat kansspelaanbieders in voorkomende gevallen een met de maatschappij c.q. wetenschap gedeeld belang hebben, bijvoorbeeld als het gaat om (onder andere) verslavingspreventie. De kanspelsector is immers ook gebaat bij het voorkomen van excessen en een positief imago. Het is ook aannemelijk dat dergelijk onderzoek in hoge mate onafhankelijk zou kunnen worden uitgevoerd, door het scherp definiëren van de rollen van en interactie tussen kansspelaanbieder en onderzoeker.

Niettemin worden er risico's gezien bij onderzoeken die door (onafhankelijke) onderzoekers worden uitgevoerd in opdracht van de kanspelsector:

- Allereerst kan er een *selectie-effect* ontstaan ten aanzien van de onderzoeken zelf: het zou zo kunnen zijn dat alleen onderzoeken waarvan verwacht wordt dat de resultaten in het belang zijn van de kanspelsector worden uitgevoerd. Aanbieders kunnen vooraf inschatten of een onderzoek waarschijnlijk gunstige of ongunstige uitkomsten voor de kanspelsector zal opleveren.
- Als het tot een onderzoek komt, kan er (bij onvoldoende waarborgen) sprake zijn van *sturing*. Een goede onafhankelijke onderzoeker zal uiteraard niet akkoord gaan met het uitvoeren van een dergelijk onderzoek. Uit de gesprekken met onderzoekers ontstaat echter het beeld dat zij de afspraken met onderzoekers die in opdracht van de kanspelsector werken niet transparant achten. Het is voor een buitenstaander daardoor moeilijk te controleren in hoeverre het onderzoek daadwerkelijk onafhankelijk is uitgevoerd. Hoewel de opdrachtgever van het onderzoek vaak een eerste indicatie geeft van de betrouwbaarheid, blijft de precieze rol van de kansspelaanbieder in het onderzoeksproces moeilijk te achterhalen.
- Tot slot is de vraag in hoeverre een kansspelaanbieder kan beslissen over *publicatie* van een onderzoek. Dit zou eveneens kunnen leiden tot een selectie-bias, waarbij alleen welgevallige resultaten zouden worden gepubliceerd.

Contact tussen onderzoeker en kansspelaanbieder

Ook wanneer onderzoekers *niet* in opdracht van kansspelaanbieders werken, is vaak contact nodig tussen onderzoeker en kansspelaanbieder (zie paragraaf 2.3 voor de huidige procedure voor het opvragen van data bij kansspelaanbieders). Vaak is daarnaast duiding nodig bij de opgeleverde data (wat overigens ook in het belang van de kansspelaanbieders is – vanuit hun perspectief is van belang dat data niet verkeerd geïnterpreteerd worden). De kansspelaanbieders kunnen, zo geven de onderzoekers aan, dit proces in sommige gevallen frustreren.

Uit de gesprekken volgt het beeld dat onderzoekers terughoudend zijn als het gaat om contact met de kansspelaanbieders. Buiten de genoemde risico's speelt hierbij de *perceptie* van niet-onafhankelijkheid een rol. Onafhankelijke onderzoekers willen ook de *schijn* voorkomen dat onderzoeken beïnvloed kunnen zijn door de kanspelsector.

Consequenties voor verbrede datadeling

Uit bovenstaande verkenning komt een aantal aanbevelingen naar voren ten aanzien van eventuele verdergaande datadeling:

- Toegang tot de data moet worden gebaseerd op het *doel* van het onderzoek, dat zuiver ten behoeve van beleidsvorming en/of wetenschap moet zijn. Hierbij moet ook worden gekeken naar mogelijke selectie-effecten ten aanzien van wel en niet uitgevoerde onderzoeken. Er zou daarom kunnen worden vereist dat de overheid, wetenschap of maatschappelijke organisaties opdrachtgever zijn, en/of worden vereist dat resultaten altijd worden gepubliceerd.
- Publicatie van de onderzoeksresultaten moet niet kunnen worden beperkt door een opdrachtgever uit de kanspelsector, om te voorkomen dat er een selectie-effect ontstaat ten aanzien van wel en niet gepubliceerde resultaten.
- Er is behoefte aan bemiddeling tussen de onderzoekers en de kansspelaanbieders bij het opvragen en duiden van data. Tegelijkertijd geven meerdere onderzoekers aan dat nauwe samenwerking met de goksector complex en onwenselijk is, omdat het de schijn van niet-onafhankelijkheid met zich meebrengt.

Een mogelijke invulling is het aanstellen van een centrale, onafhankelijke onderzoekscommissie die toeziet op wie toegang krijgt tot de data en onder welke voorwaarden. Hierbij kan worden gedacht aan een systeem vergelijkbaar met de manier waarop IVZ de toegang tot LADIS-cijfers reguleert, of de wijze waarop wordt gewerkt bij CBS Microdata. Hoewel een onafhankelijke commissie die onderzoeksplannen beoordeelt ruimte kan creëren, zijn er onder de onderzoekers ook zorgen over de werkbaarheid en de bureaucratische lasten, vooral wanneer sectorbelangen – zoals die van de kansspelindustrie – een rol zouden kunnen spelen bij de afweging van een dergelijke commissie. Zo spelen bij CBS Microdata sectorale belangen in dergelijke commissies geen rol, terwijl bij LADIS de commissie bestaat uit onder andere vertegenwoordigers uit de sector.

Volgens onderzoekers zou toegang tot spelersdata beperkt moeten blijven tot onderzoekers die verbonden zijn aan geaccrediteerde instellingen (zoals universiteiten, hogescholen, of onderzoeksinstituten) en onderzoekers met aantoonbare expertise en kennis van de AVG, bijvoorbeeld via een BROK-certificering zoals die geldt voor klinisch onderzoek. Het accreditatiesysteem van het Centraal Bureau voor de Statistiek (CBS) voor toegang tot microdata zou hier als voorbeeld kunnen dienen. In dit systeem worden zowel de onderzoeker als het onderzoeksbureau getoetst.

2.3 Huidige invulling en knelpunten

In de Wet op de kansspelen (Wok [Wok]) worden vergunninghouders (in artikel 31m lid 5) verplicht om gegevens en analyses ten aanzien van speelgedrag *geanonimiseerd* beschikbaar te stellen voor onderzoek naar kansspelverslaving. Onderzoekers kunnen op grond van dit wetsartikel de aanbieders verzoeken om data. Deze verplichting en de werkwijze zijn uitgewerkt in artikel 17 van het Besluit werving, reclame en verslavingspreventie kansspelen (Besluit Wrvk [Wrvk]) en de Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving.

Wanneer we kijken naar de mate waarin deze huidige route voorziet in de boven beschreven behoefte, komen we tot een aantal knelpunten. Het oplossen van de genoemde knelpunten kan worden gezien als *doelstellingen* voor oplossingsrichtingen. We hebben deze knelpunten voorzien van letters (A, D, K, O en P) om er verderop bij de oplossingsrichtingen aan te kunnen refereren.

Het proces voor datatoegang is suboptimaal ingericht (P)

Voor het opvragen van gegevens op basis van artikel 31m Wok is door de Ksa in overleg met de sector een procedure ingericht. Hierbij krijgen onderzoekers twee maal per jaar de kans om een aanvraag in te dienen bij aanbieders. [Ksa25] Voor het indienen van een aanvraag is een formulier opgesteld, dat beschikbaar is via de website van de Ksa. [Kan22]

Een aanvraag moet bilateraal worden ingediend (door een onderzoeker bij een specifieke aanbieder). De Ksa ontvangt daarnaast een afschrift van de verzoeken. Hoewel exacte gegevens ontbreken bij de Ksa, is het beeld op basis van gesprekken met de Ksa, aanbieders en onderzoekers, dat de hierboven beschreven procedure nog slechts zeer beperkt wordt gebruikt. Op moment van schrijven heeft één onderzoeker bij alle vergunde aanbieders een verzoek ingediend met betrekking tot de (volledige, geanonimiseerde) CDB-data.

De huidige procedure wordt door zowel onderzoekers als door de aanbieders als onpraktisch gezien.

Onderzoekers geven aan dat het feit dat slechts twee keer per jaar een aanvraag kan worden gedaan, onnodig beperkend is. De afspraak voorkomt dat aanbieders het hele

jaar door aanvragen kunnen ontvangen, maar gezien het (tot nu toe) lage aantal aanvragen lijkt dit niet een reëel probleem.

Onhandig is daarnaast dat onderzoekers een aanvraag bij iedere aanbieder afzonderlijk moeten indienen. Dit kost niet alleen meer werk, maar leidt er ook toe dat er per aanbieder discussie kan ontstaan over de aanvraag. Omdat aanbieders zelf moeten bepalen wat zij mogen en moeten delen, kan het zo zijn dat iedere aanbieder een andere interpretatie hanteert van begrippen als 'geanonimiseerd' en 'kansspelverslaving'. Het is uiteraard denkbaar dat aanbieders, naarmate er meer aanvragen worden gedaan, op deze aspecten convergeren. Daarnaast is er geen uniforme controle van de datakwaliteit, zoals die er wel is voor de CDB-gegevens.

Zoals eerder benoemd is de relatie tussen onderzoekers en aanbieders daarnaast complex – enerzijds is nauwe inhoudelijke samenwerking nodig, anderzijds is nauwe samenwerking met aanbieders voor onderzoekers lastig vanwege de (schijn van) niet-onafhankelijkheid die het zou opleveren.

Onderzoekers geven tot slot aan dat niet alle aanbieders even bereidwillig zijn om mee te werken, ondanks hun wettelijke plicht. Een onderzoeker die de procedure succesvol heeft doorlopen, merkt op dat het proces (alleen) werkbaar is dankzij de wettelijke *verplichting* voor aanbieders om gegevens aan te leveren. Dit wijst erop dat het proces zonder de wettelijke verplichting moeilijk uitvoerbaar zou zijn, omdat aanbieders dan niet op vrijwillige basis zouden willen meewerken.

Vanuit het perspectief van de *aanbieders* is de huidige procedure eveneens niet ideaal. Ieder verzoek van onderzoekers moet afzonderlijk worden behandeld, waarbij het afstemmen van de vraag, het maken van de export en het afhandelen van eventuele vervolgvragen bij iedere aanvraag opnieuw personele inzet vraagt. Vanuit de aanbieders bezien zou het prettiger zijn wanneer de datavraag van de onderzoekers in hogere mate gestandaardiseerd zou zijn.

Koppelbaarheid van spelersgegevens tussen aanbieders is niet mogelijk (K)

Artikel 31m vereist dat aanbieders de verstrekte gegevens *anonimiseren*. Doordat zij dit onafhankelijk van elkaar doen is het per definitie niet mogelijk om gegevens over het speelgedrag van een speler bij de ene aanbieder te koppelen aan gegevens over speelgedrag van diezelfde speler bij een andere aanbieder. Hierdoor is het niet mogelijk om onderzoek uit te voeren waarin het gedrag van de speler *in samenhang* wordt bekeken. Dit is echter wel zeer relevant in het kader van (onder andere) onderzoek naar verslavingspreventie. De reden hiervoor is dat spelers die door een aanbieder worden gewezen op (mogelijk) onmatig speelgedrag, en/of aan wie (vervolgens) speelbeperkingen worden opgelegd door de aanbieder, naar verwachting het spelen zullen voortzetten bij een andere aanbieder. Het gedrag van een speler kan alleen integraal worden geanalyseerd (bijvoorbeeld: het bepalen van de totale inzet van een speler bij

alle aanbieders opgeteld) wanneer ook bij een dergelijke 'switch' de spelersgegevens koppelbaar blijven.

Koppelbaarheid met achtergrondkenmerken is niet mogelijk (A)

Een ander gevolg van het vereiste om de op basis van artikel 31m gedeelde spelersgegevens te anonimiseren, is dat er ook geen ándere datasets gekoppeld kunnen worden op spelersniveau. Voor sommige onderzoeken is het echter heel relevant om bepaalde achtergrondkenmerken van spelers te kunnen meenemen in een analyse. Voorbeelden hiervan zijn leeftijd, genoten opleidingsniveau, en sociaaleconomische status. In de Microdata-omgeving van het CBS zijn deze en veel meer andere relevante gegevens beschikbaar (koppelbaar op basis van een door het CBS gedefinieerd pseudoniem, de RINpersoon, dat afgeleid is van het BSN). Een andere relevante databron zouden gegevens van banken kunnen zijn. De aanbieders beschikken over het IBAN van spelers, en banken zouden daaraan gegevens kunnen koppelen die bijvoorbeeld iets zouden kunnen zeggen over speelgedrag op *illegale* gokplatforms.

Wanneer geen achtergrondkenmerken kunnen worden gekoppeld, kunnen onderzoekers voor een deel van de kenmerken terugvallen op *proxies*. Dit zijn variabelen in de data die indirect iets zeggen over de gewenste achtergrondkenmerken. Een voorbeeld is de postcode van een speler. Op basis daarvan kan worden opgezocht in welke wijk een speler woont, en wat de gemiddelde leeftijd, opleidingsniveau en sociaaleconomische status zijn.

Onderzoek door de Ksa zélf is beperkt mogelijk (O)

De Ksa voert op dit moment met enige regelmaat onderzoek uit, en heeft hiervoor zelf onderzoekers in dienst. Als toezichthouder publiceert de Ksa periodiek een monitoringsrapportage rondom online kansspelen. [Kan25] De Ksa voert ook cross-sectioneel onderzoek uit, waarvan het recent uitgevoerde 'markers of risk'-onderzoek een voorbeeld is. [Kan251]

De Ksa zou ook *zelf* graag meer en breder onderzoek kunnen (en willen) uitvoeren, maar de mogelijkheden hiertoe zijn op dit moment beperkt. Het uitvoeren van onderzoek ten behoeve van wetenschap of beleid is op dit moment namelijk géén wettelijke taak van de Ksa. [Kan24] De Ksa kan wel onderzoek uitvoeren *ten behoeve van* haar wettelijke taken (zoals die om *risicogestuurd toezicht* te houden op de kansspelmarkt). Hierbij kan het gaan om onderzoeken die ook vanuit wetenschappelijk of beleidsperspectief relevant zijn, maar niet om onderzoeken die dat uitsluitend zijn (en niet samenhangen met een van de wettelijke taken).

Toegang tot data voor andere doelen dan kansspelverslaving is niet mogelijk (D)

Artikel 31m heeft uitsluitend betrekking op onderzoek ten behoeve van preventie van kansspelverslaving. Daardoor zijn de gegevens niet altijd bruikbaar voor onderzoeken naar andere aspecten van kansspelen zoals het bredere begrip *gokschade*, en

toekomstige beleidsdoelen daaromtrent. Zowel vanuit de aanleiding voor dit onderzoek als vanuit de onderzoekers wordt duidelijk dat deze bredere vraag bestaat.

Aanwijzingen voor toekomstige oplossingen

Tijdens de gesprekken met de onderzoekers is ook een aantal eigenschappen van mogelijke *toekomstige* oplossingsrichtingen besproken, waarbij de vraag was of, c.q. in hoeverre, deze een knelpunt zouden vormen voor de invulling van de behoefte. Hieronder volgen de belangrijkste aanwijzingen voor toekomstige oplossingsrichtingen vanuit perspectief van de onderzoekers.

Selectie-effect door toestemming

Onderzoekers geven unaniem aan dat het vragen van expliciete toestemming aan spelers voor het gebruik van hun data met terugwerkende kracht zou leiden tot een onwerkbaar selectie-effect. Hoewel 'toestemming van het subject' een AVG-grondslag is voor de verwerking van persoonsgegevens, wordt deze in de context van wetenschappelijk onderzoek als ongeschikt beschouwd.

Controle voor export van resultaten

Uit de gesprekken met onderzoekers blijkt dat het vereiste om te exporteren data voor te leggen aan een exportcommissie ter controle op celvulling (zie 3.3.2) werkbaar is. Deze inspanning zou echter wel in verhouding moeten staan tot het type onderzoek, het aantal datapunten en de benodigde hoeveelheid gegevens. De onderzoekers benadrukken dat té bureaucratische procedures relevant onderzoek onnodig kunnen belemmeren of vertragen.

De onderzoekers zien het belang van het stellen van duidelijke eisen aan de wijze waarop gegevens mogen worden gebruikt en verwerkt. Het beoordelen van individuele aanvragen kan echter veel tijd vragen, wat ten koste zou kunnen gaan van de toegankelijkheid van de data voor onderzoekers. Zo is de gewenste doorlooptijd tussen het indienen van een aanvraag en het verkrijgen van toegang tot de data, evenals de doorlooptijd tussen het indienen van de export-aanvraag en het verkrijgen van een export, idealiter zo kort mogelijk. Een van de respondenten geeft aan dat een termijn van vier tot twaalf weken redelijk zou kunnen zijn voor de totale doorlooptijd van de periode van het indienen van de aanvraag en het vrijgeven van een export. Deze doorlooptijd komt overeen met de doorlooptijd van onder andere CBS Microdata, LADIS en Translink, waarbij een gemiddelde doorlooptijd van vier tot tien weken geldt.

Werken in een externe, afgeschermdde omgeving

De gesproken onderzoekers hebben over het algemeen geen bezwaar tegen het werken in een alternatieve of afgeschermdde infrastructuur voor dataverwerking. Zij zijn gewend aan dit soort werkwijzen en accepteren eventuele extra uitvoeringsfrictie die dat oplevert, mits het de voortgang van het onderzoek niet aanzienlijk belemmert.

Publicatieplicht

Verder wordt het verplicht publiceren van onderzoeksresultaten, bijvoorbeeld via open access-kanalen, over het algemeen als acceptabel gezien. Wel bestaat er terughoudendheid ten aanzien van het opleggen van harde publicatiedeadlines, zoals een verplichting tot publicatie binnen een vast aantal maanden na dataverstrekking. Dergelijke verplichtingen kunnen onbedoeld ten koste gaan van de kwaliteit van het onderzoek. Een middenweg is wellicht om alleen te verplichten dat de resultaten van de analyse van de data binnen korte tijd openbaar worden gemaakt, en iets meer tijd te geven voor publicatie van het bijbehorende onderzoeksrapport.

Kosten voor toegang tot de data

Tot slot geven onderzoekers aan dat het vragen van een vergoeding voor toegang tot de data mogelijk een drempel opwerpt. Voorkomen moet in ieder geval worden dat een vergoeding leidt tot een selectie-effect ten aanzien van welke onderzoeken wel of niet worden uitgevoerd.

3 Juridische analyse

In dit hoofdstuk analyseren we de *juridische* mogelijkheden en obstakels om bepaalde categorieën beschikbare spelersgegevens te (her)gebruiken voor onderzoek. Dit hoofdstuk geeft antwoorden met betrekking tot de juridische component van onderzoeksvragen 2 (huidige mogelijkheden) en 3 (huidige obstakels), en analyseert concrete juridische oplossingen ten behoeve van het uitbreiden van de toegang tot data (onderzoeksvraag 4).

3.1 Juridische probleemstelling

Vanuit juridisch perspectief laten de ontwerpdoelen zoals besproken in hoofdstuk 2 zich vertalen naar de volgende vraagstukken:

1. **Rechtmatigheid.** Hoe kunnen de betreffende partijen de gegevens rechtmatig delen c.q. verwerken, daarbij in aanmerking nemend dat het (waarschijnlijk) gaat om (bijzondere) persoonsgegevens en dat daarnaast mogelijk kansspel-specifieke regelgeving van toepassing is?
2. **Medewerking.** Hoe wordt geregeld/geborgd dat de aanbieders van online kansspelen, en eventuele andere betrokken partijen, meewerken aan het delen van gegevens over speelgedrag?
3. **Taakstelling van de Ksa.** Wat is in wet- en regelgeving nodig om de Ksa breder onderzoek te kunnen laten uitvoeren, en/of datadeling te laten faciliteren?

Voor alle vraagstukken moet bepaald worden of en zo ja *hoe* dit passend kan worden gemaakt binnen de wet- en regelgeving die daarop van toepassing is. Bij het vraagstuk rondom rechtmatigheid is de vraag in hoeverre, en in welke vormen, gegevensdeling past binnen de kaders hiervoor (primair de Wet op de kansspelen en de wet- en regelgeving rondom verwerking van persoonsgegevens). Bij het medewerkingsvraagstuk is de vraag of de bestaande wet- en regelgeving voldoende mogelijkheden biedt om de medewerking te realiseren. Bij alle aspecten kan de conclusie zijn dat (gegeven een bepaalde uitwerking) *niet* kan worden geconformeerd aan bestaande wet- of regelgeving, c.q. dat de bestaande wet- en regelgeving onvoldoende houvast biedt voor het regelen/borgen van medewerking, en er nieuwe wet- of regelgeving nodig is.

Om deze vragen nader te kunnen uitwerken en beantwoorden, is het nodig om de gewenste datadeling eerst nader te duiden en af te bakenen in juridische termen. Vervolgens benoemen we een aantal uitgangspunten die we hanteren in de analyse.

3.1.1 Doel

Zoals in het vorige hoofdstuk beschreven gaat het in deze analyse om drie categorieën spelersgegevens, die aanbieders gehouden zijn te verzamelen:

- Gegevens die moeten worden verzameld ter signalering van onmatig speelgedrag op grond van art. 31m Wok en art. 13-14 Besluit Wrvk;
- Gegevens die moeten worden vastgelegd in de CDB op grond van art. 5.3 Besluit Koa;
- Gegevens die worden verzameld ter identificatie van de speler op grond van art. 4.11 Besluit Koa. (Alleen om spelers over aanbieders heen te kunnen koppelen.)

We merken hierbij op dat de gegevens in de CDB al gestandaardiseerd zijn, omdat aanbieders hun CDB moeten inrichten volgens een datamodel van de Ksa. Bij elke kansspelaanbieder zijn deze gegevens dus op dezelfde manier vastgelegd. Voor de gegevens die op grond van artikel 31m Wok worden verzameld ter signalering van onmatig speelgedrag, geldt dit echter niet. Omdat deze gegevens niet gestandaardiseerd zijn, worden ze bij verschillende aanbieders vermoedelijk op verschillende manieren bijgehouden. Potentiële oplossingen om de gegevens beschikbaar te maken voor onderzoek kunnen zijn gebaseerd op gestandaardiseerde aanlevering van de gegevens in één centrale onderzoeksdatabase, maar mogelijk ook op het faciliteren van maatwerk om de voor een onderzoek benodigde gegevens van verschillende aanbieders te ontvangen. In het volgende hoofdstuk wordt hier meer aandacht aan besteed.

We stellen niet op voorhand de voorwaarde dat de gedeelde gegevens niet te herleiden zijn tot identificeerbare personen. Er zijn namelijk modellen denkbaar, zoals bijvoorbeeld in de CBS Microdata, waarin onderzoekers in een afgeschermd omgeving toegang hebben tot gegevens over (mogelijk) identificeerbare personen, maar waarin deze gegevens enkel geëxporteerd kunnen worden nadat zij dusdanig zijn geaggregeerd en bewerkt dat deze personen redelijkerwijs niet meer te identificeren zijn.

3.1.2 Uitgangspunten

In de juridische analyse speelt een aantal uitgangspunten een centrale rol. Hieronder werken we deze uit:

De te delen spelersgegevens zijn bijzondere persoonsgegevens

Spelersgegevens die duiden op een mogelijke (gok)verslaving kwalificeren als bijzondere persoonsgegevens, namelijk gezondheidsgegevens.^a Ook als wordt onderkend dat een belangrijk deel van de spelersgegevens niet noodzakelijkerwijs

^a Zie overweging 35 AVG: "Persoonsgegevens over gezondheid dienen alle gegevens te omvatten die betrekking hebben op de gezondheidstoestand van een betrokkene en die informatie geven over de lichamelijke of geestelijke gezondheidstoestand van de betrokkene in het verleden, het heden en de toekomst." Zie ook de MvT van de Wet Koa, p. 46: "De registratie en analyse van gegevens met betrekking tot het speelgedrag door de exploitant van speelhallen, speelcasino's en kansspelen op afstand kunnen op enig moment ook leiden tot de verwerking van bijzondere persoonsgegevens, te weten gezondheidsgegevens in de zin van artikel 16 Wbp."

gezondheidsgegevens betreft, kan niet met zekerheid worden *uitgesloten* dat in een aantal, niet op voorhand bekende gevallen uit deze gegevens is op te maken dat iemand verslaafd is. Dit betekent dat zekerheidshalve *alle* spelersgegevens moeten worden behandeld als gezondheidsgegevens.^a

Het gevolg hiervan is dat het verwerken van de spelersgegevens verboden is, op grond van artikel 9(1) AVG, tenzij een uitzondering van artikel 9(2) AVG van toepassing is.

De spelersgegevens zijn in beheer van de kansspelaanbieders

De gegevens waar het in het voorliggende vraagstuk om draait, worden verzameld door de kansspelaanbieders in het kader van aan hen opgelegde wettelijke verplichtingen. Toegang tot deze data voor onderzoeksdoeleinden vereist dan ook dat de data worden *gedeeld* door de kansspelaanbieders.

Omdat het echter bijzondere persoonsgegevens zijn (zie voorgaande), mogen zij deze in beginsel niet delen, met uitzondering van specifieke in de wet geregelde situaties. Naast de privacywetgeving en de kansspelregelgeving zelf zijn ons tijdens het onderzoek geen andere wettelijke beperkingen bekend geworden die op de aanbieders van toepassing zijn met betrekking tot het delen van de gegevens. Behalve juridische aspecten kunnen er uiteraard ook andere overwegingen spelen die zich verzetten tegen het delen van gegevens. Zo kan het vanuit oogpunt van concurrentie tussen aanbieders onwenselijk zijn als gedeelde gegevens voor andere aanbieders beschikbaar zouden komen.

De Ksa heeft toegang tot de spelersgegevens, maar kan deze niet delen

De Ksa heeft op grond van artikel 34l Wok toegang tot alle bedoelde gegevens, maar enkel voor zover dat redelijkerwijs noodzakelijk is voor de vervulling van haar (wettelijke) taken. Deze taken omvatten niet het mogelijk maken van onderzoek voor wetenschappelijke of statistische doeleinden.^b

In theorie kunnen (bijzondere) persoonsgegevens met (uitdrukkelijke) toestemming van de desbetreffende speler worden gebruikt voor wetenschappelijk onderzoek. In de praktijk kan dat echter vaak niet.

Voor het verwerken van persoonsgegevens moet er altijd ten minste één van de in de AVG genoemde grondslagen zijn. Een van deze grondslagen is *toestemming van de*

^a Merk op dat het verwijderen van alle variabelen die *mogelijk* iets zeggen over verslaving, er logischerwijs toe zou leiden dat geen onderzoek meer kan worden gedaan naar kansspelverslaving of speelgedrag. Daardoor zou niet meer in de onderzoeksbehoefte kunnen worden voorzien. Het verwijderen van een deel van de gegevens zodat enkel reguliere persoonsgegevens overblijven is dus geen optie.

^b Zie artikel 33b Wok voor een opsomming van de taken van de Ksa en [\[kansspelautoriteit.nl\]](https://kansspelautoriteit.nl) voor een toelichting per taak.

persoon op wie de persoonsgegevens betrekking hebben. Om een aantal redenen is toestemming in de voorliggende casus echter geen haalbare grondslag:

- Voor het gebruik van toestemming als grondslag is vereist dat de toestemming *vrijelijk gegeven* is. In de praktijk zullen er echter altijd twijfels kunnen zijn of de toestemming in vrijheid is gegeven, wanneer deze online wordt gevraagd aan personen die mogelijk gokverslaafd zijn.
- Een ander vereiste is dat de toestemming *specifiek* moet zijn. Aangezien de uit te voeren onderzoeken niet vooraf expliciet zijn uitgewerkt, moet echter ofwel eenmaal een 'generieke' toestemming worden gevraagd voor de verwerking ten behoeve van onderzoek, ofwel elke keer een (onderzoeks)specifieke toestemming per onderzoek. In het eerste geval wordt mogelijk niet voldaan aan het vereiste dat de toestemming specifiek is gegeven. Het tweede kent praktische bezwaren en leidt naar verwachting tot een veel lagere dekking van de data: spelers moeten voor elk onderzoek opnieuw beslissen of ze toestemming willen geven en veel spelers zullen ten tijde van de uitvraag niet inloggen en dus geen toestemming kunnen geven.
- Wanneer toestemming als grondslag wordt gebruikt, staat het de betrokkenen vrij om hun toestemming op een later moment in te trekken. Afhankelijk van de wijze waarop de data beschikbaar wordt gemaakt voor onderzoek, kan het voor praktische problemen zorgen als spelers dit doen.
- Het enkel meenemen van gegevens van spelers die toestemming hebben gegeven zal vrijwel zeker tot een *selection bias* in de data leiden: de groep subjecten die toestemming geeft is niet representatief voor de hele populatie. Dit betekent dat de data voor veel onderzoeken niet meer goed bruikbaar is. Om deze redenen is toestemming in veel gevallen géén haalbare grondslag.

3.2 Borgen van rechtmatigheid datadeling

3.2.1 Juridische obstakels

Vanuit juridisch oogpunt zijn er twee categorieën regelgeving die een obstakel kunnen vormen voor het rechtmatig beschikbaar stellen van spelersdata voor onderzoek: kansspelregelgeving en privacyregelgeving. Bij de eerste categorie ligt een mogelijk obstakel in de formulering en uitwerking van artikel 31m lid 5 Wok. Bij de tweede gaat het met name om het verbod uit de AVG op de verwerking van bijzondere persoonsgegevens. In de volgende paragrafen gaan we hier achtereenvolgens op in.

Artikel 31m lid 5 Wok

Zoals eerder benoemd bepaalt artikel 31m van de Wok dat online kansspelaanbieders gegevens over speelgedrag moeten verzamelen ter signalering van onmatig speelgedrag. Het vijfde lid stelt met betrekking tot deze gegevens:

“De vergunninghouder stelt gegevens en analyses als bedoeld in het eerste lid geanonimiseerd beschikbaar voor onderzoek naar kansspelverslaving.”

Gedurende het onderzoek hebben we gemerkt dat over de interpretatie van deze bepaling discussie mogelijk is. In de kern is de vraag of de bepaling slechts een verplichting creëert waar de kansspelaanbieders aan moeten voldoen, of dat de wetgever (ook) heeft willen beperken wat er met de gegevens mag gebeuren.

Voor de eerste interpretatie pleit de zinsbouw en de wijze waarop de bepaling is uitgewerkt in de onderliggende regelgeving, namelijk artikel 17 van het Besluit werving, reclame en verslavingspreventie kansspelen (Besluit wrvk) en de daaronder liggende Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving:

“De houder van een vergunning tot het organiseren van kansspelen op afstand verstrekt op verzoek van een onderzoeksinstelling overeenkomstig door de raad van bestuur gestelde regels geanonimiseerde gegevens als bedoeld in artikel 31m, vijfde lid, van de wet, ten behoeve van wetenschappelijk onderzoek naar kansspelverslaving.” [Wrvk art. 17]

“Deze regeling heeft uitsluitend betrekking op de verplichting van de vergunninghouder om op verzoek van een onderzoeksinstelling gegevens te verstrekken ten behoeve van wetenschappelijk onderzoek naar kansspelverslaving, als bedoeld in artikel 31m, vijfde lid, van de wet en artikel 17 van het besluit.” [Kan21 art. 2.1]

Dit lezende lijkt artikel 31m lid 5 Wok enkel bedoeld om te zorgen dat kansspelaanbieders verplicht zijn om geanonimiseerde gegevens aan te leveren als onderzoekers daarnaar vragen ten behoeve van hun onderzoek naar kansspelverslaving. Oftewel: onderzoekers die aan de voorwaarden voldoen en die geanonimiseerde gegevens willen gebruiken voor onderzoek naar kansspelverslaving, maken er op basis van deze bepaling aanspraak op dat aanbieders hen deze geanonimiseerde gegevens inderdaad verstrekken. De bepaling verplicht de aanbieders de gegevens te verstrekken.

Daarmee is evenwel *niet* noodzakelijkerwijs gezegd dat het voor aanbieders verboden is om gegevens aan te leveren voor onderzoek naar iets anders dan kansspelverslaving. De bepaling geeft onderzoekers geen aanspraak op deze gegevens, maar staat er niet aan in de weg dat ze toch een verzoek doen om verstrekking van deze gegevens. En aanbieders zijn niet verplicht om aan een dergelijk verzoek gehoor te geven, maar het is daarmee niet gezegd dat het aanbieders verboden is om niet-geanonimiseerde gegevens te verstrekken; zij zijn alleen niet *verplicht* om dit te doen (en het kan wel verboden zijn op grond van andere regelgeving, zoals de AVG). Tot slot zou het artikel er niet aan in de weg staan dat een onderzoeker de gegevens, die hij heeft ontvangen voor onderzoek naar kansspelverslaving, later óók gebruikt voor ander type kansspelonderzoek (al ligt voor de hand dat vooraf wordt vastgelegd waarvoor de gegevens zullen worden gebruikt).

De bepaling kan echter ook worden uitgelegd als een beperking op wat er met de gegevens mag gebeuren. In deze restrictieve interpretatie mogen de gegevens die aanbieders vastleggen ter signalering van onmatig speelgedrag **alleen geanonimiseerd** beschikbaar worden gesteld en **enkel voor onderzoek naar kansspelverslaving**. In deze lezing zou de bepaling het delen van niet-geanonimiseerde gegevens verbieden, zelfs wanneer de AVG er niet aan in de weg zou staan en de aanbieder de gegevens wél zou *willen* delen. Zelfs het delen van geanonimiseerde gegevens voor andere doelen dan onderzoek naar kansspelverslaving zou dan niet toegestaan zijn. Dat zou een vergaande beperking zijn.

Gelet op de zinsbouw en woordkeuze van de wetgever, en de zojuist geciteerde lagere regelgeving, is dit geen voor de hand liggende uitleg van de bepaling.^a Daartegenover staat dat de toelichting bij artikel 31m Wok in parlementaire stukken wel (enige) steun lijkt te kunnen bieden voor een restrictieve lezing, waarin de wetgever als het ware heeft willen ‘dichttimmeren’ waarvoor de gegevens mogen worden gebruikt:

“Op grond van het vijfde lid van de artikelen 27ja, 30v en 31m worden in de lagere regelgeving nadere regels gesteld met betrekking tot onder meer de verwerking van persoonsgegevens door de vergunninghouder, de waarborgen voor de rechtmatige verwerking daarvan en passende technische en organisatorische maatregelen ter beveiliging tegen verlies of onrechtmatige verwerking. Daarbij zal onder meer worden bepaald dat de vergunninghouder dergelijke gegevens niet verstrekt aan anderen dan de kansspelautoriteit voor zover deze die gegevens heeft voor de goede uitoefening van haar taken. In geanonimiseerde vorm kunnen gegevens beschikbaar worden gesteld voor wetenschappelijk onderzoek naar kansspelverslaving.”^b

Eenduidig is dit echter niet. In de eerste plaats merken we op dat deze toelichting, op de laatste zin na, feitelijk niet gaat over het vijfde lid van artikel 31m maar over het zesde lid, waarin staat dat met betrekking tot de voorgaande leden bij lagere regelgeving nadere regels worden gesteld over onder meer de registratie van de gegevens en de verwerking van persoonsgegevens.^c In de tweede plaats kan uit deze toelichting weliswaar worden opgemaakt dat de wetgever heeft willen beperken wat er met de betreffende gegevens gebeurt, meer specifiek door in de lagere regelgeving te

^a Het is moeilijk een elegante bepaling te formuleren waarin zowel de verplichting om gegevens te verstrekken ligt besloten, als de beperking aan wat er (los van die verplichting) met de gegevens mag gebeuren. Als bijvoorbeeld het woord ‘uitsluitend’ wordt toegevoegd, waarmee de beperking duidelijk wordt, dan valt de verplichting weg. Er is eigenlijk ook geen reden om deze twee zaken in één bepaling te regelen. Op basis van de lagere regelgeving is duidelijk dat met lid 5 in ieder geval is bedoeld de verplichting te creëren. De beperking zou men daarom verwachten in een ander lid of in de lagere regelgeving op basis van art. 31m lid 6 Wok, zie ook het vervolg van deze paragraaf.

^b Tweede Kamer, vergaderjaar 2013–2014, 33 996, nr. 3, p. 47 (MvT Wet Koa).

^c In de artikelen 27ja en 30v van de Wok, waarnaar ook wordt verwezen, is dit wel het vijfde lid.

verbieden dat kansspelaanbieders de gegevens verstrekken aan anderen dan de Ksa, waarop art. 31m lid 5 dan een uitzondering zou zijn, maar een dergelijk verbod vinden we in de lagere regelgeving niet terug. De toelichting in de parlementaire stukken geeft dus een intentie weer om in lagere regelgeving beperkingen op te nemen voor wat er met de gegevens kan gebeuren, maar aan die intentie is vervolgens géén gevolg gegeven.

Ook de bewoording van de slotzin van de aangehaalde toelichting – in geanonimiseerde vorm kunnen gegevens beschikbaar worden gesteld – lijkt te impliceren dat het niet mogelijk is om niet-geanonimiseerde gegevens beschikbaar te stellen. Het probleem blijft echter dat dit in de regelgeving zelf niet terug te zien is.

De in de toelichting uitgedrukte intentie van de wetgever zou aanleiding kunnen zijn voor een restrictieve interpretatie, die met zich brengt dat er op grond van de Wok geen ruimte zou zijn voor kansspelaanbieders om de betreffende gegevens te delen zonder ze te anonimiseren, of om ze te delen voor onderzoek naar een ander onderwerp dan kansspelsverslaving (zelfs wanneer ze wel geanonimiseerd zijn). Op basis van de regelgeving zelf zouden we echter beargumenteren dat deze ruimte er wel degelijk is. Van belang daarbij achten wij dat de toelichting weliswaar behulpzaam kan zijn bij de uitleg van de bepaling en inzicht kan geven in de door de wetgever gemaakte keuzes, maar niet kan worden gebruikt om extra regels te stellen die niet uit de bepaling blijken. Evenmin kan een toelichting worden gebruikt voor een nadere omlijning of beperking van in de regeling voorkomende termen of begrippen.^a Van belang achten wij verder dat geen gevolg is gegeven aan de in de toelichting benoemde intentie om in lagere regelgeving het gebruik van de gegevens te beperken. Ook dat is een aanwijzing dat dit gegevensgebruik niet noodzakelijkerwijs is beperkt tot de beschikbaarstelling van de gegevens uitsluitend in geanonimiseerde vorm en uitsluitend voor onderzoek naar kansspelsverslaving.

In het vervolg van dit rapport hanteren we daarom de, volgens ons correcte, ruimere interpretatie van de huidige regelgeving, die meebrengt dat de Wok zelf niet in de weg staat aan het delen van niet-geanonimiseerde spelersgegevens of het delen van spelersgegevens voor andere doelen dan onderzoek naar kansspelsverslaving. Dit betekent dat we voor rechtmatige gegevensverstrekking enkel nog naar de AVG hoeven te kijken. Aan het slot van dit hoofdstuk is een alinea opgenomen over de impact die het hanteren van de restrictieve interpretatie zou hebben op de in dit hoofdstuk te behandelen oplossingen.

^a Aanwijzing 4.47 Aanwijzingen voor de regelgeving; zie ook bijv. Kamerstukken I 2017/18, 34 453, p. 4

Om buiten twijfel te stellen dat:

1. de bepaling van lid 5 slechts een verplichting vormt en geen beperking vormt op het gebruik van de gegevens, en
2. het delen van niet-geanonimiseerde gegevens toegestaan is als wordt voldaan aan de voorwaarden die de AVG daaraan stelt, en
3. dat geanonimiseerde gegevens ook voor andere onderzoeksdoeleinden kunnen worden gedeeld,

zou aan artikel 31m lid 5 Wok wellicht een zin toegevoegd kunnen worden:

De vergunninghouder stelt gegevens en analyses als bedoeld in het eerste lid geanonimiseerd beschikbaar voor onderzoek naar kansspelverslaving. De verplichting om hieraan mee te werken laat onverlet dat deze gegevens en analyses ook niet-geanonimiseerd of voor andere onderzoeksdoeleinden kunnen worden gedeeld, mits gebruik wordt gemaakt van passende waarborgen om de belangen van betrokkenen te beschermen.

AVG

Naast de eventuele beperking vanuit de Wok zelf, vormt het verbod op het verwerken van bijzondere persoonsgegevens (art. 9 lid 1 AVG) de belangrijkste juridische beperking voor de voorziene datadeling. Juridisch zijn er twee voor de hand liggende routes om hiermee om te gaan:

1. **Zorgen dat er géén persoonsgegevens worden verwerkt (anonimiseren).** Een van de routes is te zorgen dat er überhaupt geen persoonsgegevens worden verwerkt, door uit te sluiten dat de identiteit van de desbetreffende spelers redelijkerwijs kan worden achterhaald door degenen met toegang tot de data. Dit kan worden bereikt door dusdanig solide te pseudonimiseren dat de betrokkenen niet te (her)identificeren zijn, ook niet wanneer degenen met toegang tot de data over aanvullende (koppelbare) informatie beschikken. In dat geval is feitelijk sprake van anonimisering.
2. **Gebruik maken van de uitzondering voor onderzoeksdoeleinden.** De tweede route is om gebruik te maken van artikel 24 UAVG. [UAVG] Dit artikel biedt onder bepaalde voorwaarden een uitzondering op het verbod om bijzondere persoonsgegevens te verwerken, voor wanneer de verwerking noodzakelijk is voor wetenschappelijk of statistisch onderzoek. Naast deze uitzondering is ook een grondslag voor de verwerking vereist.

In de volgende paragrafen werken we de beide routes nader uit.

3.2.2 Oplossingen voor de juridische obstakels

In deze paragraaf bespreken we de juridische routes om de spelersgegevens op rechtmatige wijze te verwerken en delen, zodat ze ter beschikking kunnen worden gesteld aan (uiteindelijk) onderzoekers.

Op wie/wat zijn deze routes van toepassing?

Een belangrijke vraag hierbij is op welke entiteit(en) deze routes, en de bijbehorende voorwaarden, van toepassing zijn. Wie deelt de gegevens met wie, en welke partij is verwerkingsverantwoordelijke voor welke verwerking? Het antwoord op deze vragen is voor een deel afhankelijk van de te kiezen oplossingsrichting:

- De aanbieders moeten de gegevens *mogen* delen met de onderzoeker, en/of (als daar sprake van is) een tussenpartij (vrijwillig of verplicht).
- Een tussenpartij (als daar sprake van is) moet de gegevens *mogen* verwerken en delen met onderzoekers.
- De onderzoekers moeten de gegevens *mogen* verwerken en moeten hun resultaten *mogen* publiceren.

Voor de datadeling zijn verschillende oplossingsrichtingen denkbaar, waarbij onder andere tussenpartijen een rol kunnen spelen. De juridische analyse in dit hoofdstuk is niet specifiek toegespitst op een concrete richting, maar gaat abstracter in op de juridische omstandigheden waaronder spelersgegevens gedeeld mogen worden.

Route R1: zorgen dat geen sprake is van persoonsgegevens

Het juridisch kader voor het gebruik van persoonsgegevens wordt in belangrijke mate bepaald door gegevensbeschermingswetgeving, meer specifiek de AVG. Deze regels zijn van toepassing op de verwerking van persoonsgegevens, dat wil zeggen gegevens over geïdentificeerde of identificeerbare natuurlijke personen.

Voor de mogelijkheden en onmogelijkheden om gebruik te maken van spelersgegevens is daarom vooral van belang in hoeverre deze gegevens kwalificeren als persoonsgegevens. Of, anders gezegd, in hoeverre de spelers op wie de gegevens betrekking hebben (subjecten) identificeerbaar zijn.

Als identificeerbaar wordt op grond van de AVG beschouwd: een persoon die direct of indirect kan worden geïdentificeerd, met name aan de hand van een zogenoemde 'identificator', zoals een naam, een identificatienummer, locatiegegevens of een online identificator, of aan de hand van een of meer elementen die kenmerkend zijn voor de

fysieke, fysiologische, genetische, psychische, economische, culturele of sociale identiteit van de persoon.^a

Om te bepalen in hoeverre er sprake is van identificeerbaarheid, wordt gekeken naar alle middelen waarvan redelijkerwijs valt te verwachten dat zij worden gebruikt door de verwerkingsverantwoordelijke of door een andere persoon om de identiteit vast te stellen. Om uit te maken of van middelen redelijkerwijs valt te verwachten dat zij zullen worden gebruikt, moet rekening worden gehouden met alle objectieve factoren, zoals de kosten van en de tijd benodigd voor identificatie, met inachtneming van de beschikbare technologie op het tijdstip van verwerking en de technologische ontwikkelingen.^b Als een onevenredige hoeveelheid tijd, geld en menskracht nodig zou zijn om identificatie te realiseren, is er geen sprake van persoonsgegevens.

Het voorgaande biedt ruimte voor de mogelijkheid dat spelersgegevens op zodanige wijze worden gepseudonimiseerd, en eventueel bewerkt en de verwerking ervan met wettelijke waarborgen omkleed, dat deze voor andere partijen niet (meer) kwalificeren als persoonsgegevens. Dat deze ruimte er inderdaad is, is in een recent arrest van het Hof van Justitie van de EU expliciet bevestigd:

"It follows that [...] pseudonymised data must not be regarded as constituting, in all cases and for every person, personal data [...], in so far as pseudonymisation may, depending on the circumstances of the case, effectively prevent persons other than the controller from identifying the data subject in such a way that, for them, the data subject is not or is no longer identifiable."^{cd}

Voorafgaand aan dit arrest was hier nogal eens discussie over, waarbij sommige partijen, zoals de Europese toezichthouder voor gegevensbescherming, het standpunt

^a Artikel 4(1) AVG.

^b Overweging 26 van de AVG. [33 p. 6.2.4].

^c HvJEU 4 september 2025, C-413/23 P, ECLI:EU:C:2025:645 (*GAR/EDPS*), r.o. 86.

^d We nemen hier het Engelse citaat op omdat er tijdens het schrijven van dit rapport nog geen definitieve Nederlandse vertaling is gepubliceerd en omdat de gepubliceerde voorlopige Nederlandse vertaling een fout lijkt te bevatten:

"Hieruit volgt dat, anders dan de EDPS stelt, gepseudonimiseerde gegevens niet in alle gevallen en voor eenieder mogen worden beschouwd als persoonsgegevens [...], aangezien pseudonimisering, afhankelijk van de omstandigheden van het geval, daadwerkelijk kan verhinderen dat andere personen dan de verwerkingsverantwoordelijke de betrokkene op zodanige wijze identificeren dat deze betrokkene voor hen niet of niet meer identificeerbaar is."

De gedachte dat iemand een betrokkene op zodanige wijze kan identificeren dat deze betrokkene niet meer identificeerbaar is, is inherent tegenstrijdig. Het zinsdeel 'op zodanige wijze' heeft in de vertaling betrekking gekregen op 'identificeren', maar zou betrekking moeten hebben op 'verhinderen'. In de Engelse versie is dit taalkundig ambigu, maar de enige logische interpretatie is dat 'in such a way' betrekking heeft op 'prevent', niet op 'identifying'.

innamen dat gepseudonimiseerde gegevens altijd persoonsgegevens waren. Zie ter illustratie van dit debat een korte bespreking van *GAR/EDPS* hieronder.

GAR/EDPS

De vraag of gepseudonimiseerde gegevens per definitie persoonsgegevens zijn, kwam recent uitgebreid aan bod in de zaak *GAR/EDPS* (Engelse versie: *SRB/EDPS*).^a Bij het bepalen of de in deze zaak overgedragen gepseudonimiseerde gegevens al dan niet moeten worden aangemerkt als persoonsgegevens, past het Gerecht de zogenaamde **relatieve benadering** toe. Dat wil zeggen dat vanuit het perspectief van de *ontvanger* van de gepseudonimiseerde gegevens moet worden bekeken of sprake is van persoonsgegevens. Dit betekent niet dat gepseudonimiseerde gegevens geen persoonsgegevens zijn zolang de ontvangende partij niet over de sleutel beschikt, maar wel dat vanuit de positie van de ontvanger moet worden beoordeeld of er sprake is van middelen waarvan redelijkerwijs is te verwachten dat die worden gebruikt om de desbetreffende personen te identificeren. Het Gerecht past hiermee de relatieve benadering uit het *Breyer*-arrest expliciet toe op een overdracht van gepseudonimiseerde gegevens.

De European Data Protection Supervisor (EDPS) heeft tegen deze uitspraak beroep aangetekend, waarin het onder meer betoogt dat de zojuist beschreven relatieve benadering in de betreffende zaak niet passend is. De EDPS pleit in feite voor een **absolute benadering**: omdat de informatie aan de hand waarvan de betrokkenen geïdentificeerd *kunnen* worden (de sleutel) blijft bestaan, blijven de betrokkenen identificeerbaar en blijven de gepseudonimiseerde gegevens dus persoonsgegevens.

A-G Spielmann beargumenteert in zijn conclusie voor het Hof van Justitie echter dat de AVG de mogelijkheid open laat dat betrokkenen na pseudonimisering niet identificeerbaar zijn (punten 46-52), dat dit niet in strijd is met het doel om een hoog niveau van bescherming van persoonsgegevens te waarborgen (punten 53-58) en dat men in de betreffende zaak had moeten nagaan *“of de pseudonimisering van de betrokken gegevens voldoende solide was om tot de slotsom te komen dat de [betrokkenen] [...] redelijkerwijs niet identificeerbaar waren. Met andere woorden, als [ontvanger] in deze context over redelijke middelen beschikte om de [betrokkenen] te identificeren, kon zij geacht worden persoonsgegevens te verwerken.”* (punt 59).^b

De A-G is dus, net als het Gerecht, van oordeel dat gepseudonimiseerde gegevens niet per definitie persoonsgegevens zijn, enkel omdat *ergens* een sleutel beschikbaar is en er een theoretische mogelijkheid tot identificatie is. In lijn met *Breyer* en enkele

^a Gerecht EU 26 april 2023, T-557/20, ECLI:EU:T:2023:219 (*GAR/EDPS*).

^b Conclusie van A-G D. SPIELMANN van 6 februari 2025 bij zaak C-413/23 P (*GAR/EDPS*).

andere arresten van het Hof van Justitie stelt zowel het Gerecht als de A-G dat bepalend is of er *redelijke middelen* zijn om de betrokkenen te identificeren, en dat dit *vanuit de positie van de ontvanger* moet worden bekeken om te bepalen of de gegevens *voor de ontvanger* persoonsgegevens zijn.

Het Hof van Justitie volgt de AG en het Gerecht:

“It follows that, contrary to what the EDPS maintains, pseudonymised data must not be regarded as constituting, in all cases and for every person, personal data for the purposes of the application of Regulation 2018/1725, in so far as pseudonymisation may, depending on the circumstances of the case, effectively prevent persons other than the controller from identifying the data subject in such a way that, for them, the data subject is not or is no longer identifiable.”^a

Gepseudonimiseerde gegevens zijn dus niet per definitie persoonsgegevens. Voor onze casus over spelersgegevens betekent dit dat er, wat gegevensbeschermingsrecht betreft, ruimte is voor een constructie waarin kansspelaanbieders de betreffende spelersgegevens pseudonimiseren op zodanige wijze dat de gegevens wel koppelbaar zijn over aanbieders heen, maar voor de ontvangers (de onderzoekers of eventuele tussenpartijen) niet kwalificeren als persoonsgegevens. Voorwaarde voor zo'n constructie is dat de ontvangers niet beschikken over middelen waarvan redelijkerwijs valt te verwachten dat die gaan worden gebruikt om de spelers te identificeren.

Over de vraag of een theoretische mogelijkheid tot heridentificatie *“een middel vormt waarvan mag worden aangenomen dat het redelijkerwijs kan worden ingezet om de betrokken persoon te identificeren”*, stelt het Hof van Justitie in *Breyer* dat dit niet het geval is *“indien de identificatie van de betrokkene bij de wet verboden wordt of in de praktijk ondoenlijk is, bijvoorbeeld omdat zij – gelet op de vereiste tijd, kosten en mankracht – een excessieve inspanning vergt, zodat het gevaar voor identificatie in werkelijkheid onbeduidend lijkt.”^b* Hiervoor kunnen zowel technische als juridische maatregelen worden gebruikt. Hieronder gaan we daar nader op in.

Ook ná publicatie moeten de rechten van spelers uiteraard worden beschermd en moeten spelers niet-identificeerbaar blijven. Waar aan de beheerder van de database (de zgn. tussenorganisatie) en aan mensen die toegang krijgen tot de gegevens nog allerlei wettelijke eisen gesteld kunnen worden, kan dit voor het brede publiek dat een openbare publicatie te zien krijgt uiteraard niet. Het is dan mogelijk dat er mensen zijn die over extra informatie beschikken waardoor zij betrokkenen redelijkerwijs zouden

^a HvJEU 4 september 2025, C-413/23 P, ECLI:EU:C:2025:645 (*GAR/EDPS*), r.o. 86.

^b HvJEU 19 oktober 2016, C-582/14, ECLI:EU:C:2016:779 (*Breyer*), r.o. 45-46.

kunnen identificeren. Uit het *Olaf*-arrest blijkt dat hiermee rekening moet worden gehouden en dat gegevens hierdoor als persoonsgegevens kunnen worden aangemerkt, ook als identificatie door het grote publiek praktisch niet haalbaar is.^a Een juridische maatregel die hier wellicht voor de hand ligt betreft het bij wet of desnoods contractueel aan de onderzoekers opleggen van een verplichting om ervoor zorg te dragen dat hun publicaties niet (redelijkerwijs) kunnen leiden tot identificatie van de desbetreffende spelers.^b

Redelijke middelen om te identificeren

Het gaat er dus om dat er maatregelen worden genomen, en waarborgen worden ingesteld, die voorkomen dat er middelen zijn waarvan redelijkerwijs is te verwachten dat zij worden gebruikt om de identiteit van de spelers te achterhalen. Er valt daarbij allereerst te denken aan technische maatregelen, maar in aanvulling daarop kan ook worden gedacht aan juridische of organisatorische maatregelen.

Voordat we ingaan op mogelijke maatregelen, staan we stil bij mogelijke middelen om betrokkenen te identificeren. Deze zijn in te delen in twee categorieën: identificatie door koppeling met andere gegevens (bijvoorbeeld de sleutel) en identificatie op basis van de gegevens zelf.

Koppeling met andere gegevens

De meest voor de hand liggende methode om in het geval van gepseudonimiseerde gegevens de betrokkenen te identificeren, is het terugdraaien van de pseudonimisering aan de hand van de sleutel. Dit kan in principe worden gedaan door iedereen die toegang heeft tot zowel de gepseudonimiseerde gegevens als de sleutel.^c

Hoewel het ongedaan maken van pseudonimisering door gebruik van de daarvoor bedoelde sleutel gevoelsmatig wellicht anders is dan het identificeren van betrokkenen

^a HvJ EU 7 maart 2024, C-479/22 P, ECLI:EU:C:2024:215 (*Olaf*), r.o. 60-61. In casu ging het om personen die op hetzelfde wetenschappelijke gebied werkzaam waren en die de beroepsloopbaan van de betrokkene kenden, waardoor zij deze zonder excessieve inspanning zouden kunnen identificeren.

^b Zie bijvoorbeeld Kst II 2020/21, 35479, nr. 3, p. 6-7, over een vergelijkbaar verbod bij het voorgenomen gebruik van telecom-locatiegegevens t.b.v. het bestrijden van COVID-19: “Hoewel er studies zijn die de – theoretische – mogelijkheid aantonen om in bepaalde gevallen ook geaggregeerde locatiedata te herleiden tot identificeerbare natuurlijke personen, valt niet redelijkerwijs te verwachten dat de verwerkingsverantwoordelijke of een andere persoon deze middelen gebruikt. De kosten van en de tijd benodigd voor identificatie zonder daarbij gebruik te kunnen maken van de brongegevens (de locatie- en verkeersgegevens die de aanbieders beheren) maken een dergelijke identificatie onwaarschijnlijk. Belangrijk hierbij is dat het de aanbieders op grond van de Telecommunicatiewet verboden is om de brongegevens aan derden ter beschikking te stellen.”

^c Die sleutel kan bijvoorbeeld een koppeltabel zijn, waarin de pseudoniemen worden gekoppeld aan een directe identicator. Andere typen sleutels zijn ook mogelijk. Het gaat erom dat er ergens informatie beschikbaar is waarmee de pseudonimisering kan worden teruggedraaid en betrokkenen kunnen worden geïdentificeerd.

door de gegevens te koppelen aan een compleet andere dataset, komt dit in wezen op hetzelfde neer. In beide gevallen is er een dataset met betrokkenen die niet direct identificeerbaar zijn, maar door koppeling met andere gegevens wel. Hieronder tonen we enkele relevante voorbeelden om te illustreren wat in dit soort situaties als redelijk middel om betrokkenen te identificeren wordt beschouwd.

- *Scarlet Extended*: Het ging in dit arrest om internetgebruikers die aan de hand van hun IP-adressen door hun internetprovider geïdentificeerd konden worden. Van belang is dat de IP-adressen werden verzameld *door de internetproviders die zelf ook beschikten over de abonneegegevens die nodig waren om de betrokkenen aan de hand van hun IP-adres te identificeren*. In deze situatie was het Hof van oordeel dat de gegevens voor de internetproviders kwalificeerden als persoonsgegevens.^a
- *Breyer*: In dit arrest ging het wederom om het identificeren van personen aan de hand van IP-adressen, maar de partij die hier de IP-adressen registreerde was niet de internetprovider maar een websiteaanbieder. Deze beschikte niet zelf over de benodigde extra informatie (abonneegegevens) om de internetgebruikers te identificeren. Het Hof stelde dat, om te kunnen spreken van persoonsgegevens, niet vereist is dat alle voor identificatie benodigde informatie bij dezelfde partij berust. Waar het om gaat is of de informatie te combineren is met de benodigde extra informatie. Omdat de websiteaanbieder in dit geval over juridische mogelijkheden beschikte om de benodigde informatie van de internetprovider te verkrijgen, oordeelde het Hof dat de websiteaanbieder de betrokkenen kon identificeren en dat de IP-adressen voor deze websiteaanbieder dus kwalificeerden als persoonsgegevens.^b
- *Kentekensparkeren*: In een arrest van het Gerechtshof Amsterdam was de vraag aan de orde of, en voor wie, (versleutelde) kentekengegevens persoonsgegevens zijn. Bij het parkeren voert de bestuurder het kenteken in, wat versleuteld in een centrale opslag wordt opgeslagen. Bij controle door de heffingsambtenaar worden gescande kentekens eveneens versleuteld naar de centrale opslag gestuurd om te controleren of het betreffende kenteken daar is aangemeld. Het Gerechtshof oordeelde dat de versleutelde kentekengegevens in de centrale opslag geen persoonsgegevens zijn, omdat deze niet tot identificeerbare natuurlijke personen herleidbaar zijn, maar dat een kenteken voor de heffingsambtenaar wel een persoonsgegeven is. Dit laatste omdat de

^a HvJ EU 24 november 2011, C-70/10, ECLI:EU:C:2011:771 (*Scarlet Extended*), r.o. 51 en daarover HvJ EU 19 oktober 2016, C-582/14, ECLI:EU:C:2016:779 (*Breyer*), r.o. 33-34.

^b HvJ EU 19 oktober 2016, C-582/14, ECLI:EU:C:2016:779 (*Breyer*), r.o. 38-49. Dit oordeel is herhaald in het arrest *Mircom*, waar een vergelijkbare situatie speelde, zie HvJ EU 17 juni 2021, C-597/19, ECLI:EU:C:2021:492 (*Mircom*), r.o. 102.

heffingsambtenaar een kenteken kan relateren aan een identificeerbaar persoon na gegevensverstrekking door de RDW.^a

- *NZa HoNOS+*: In deze uitspraak van de rechtbank Midden-Nederland werden door behandelaren in de GGZ gegevens over de geestelijke gezondheid van hun cliënten (HoNOS+-gegevens) verstrekt aan de NZa. Het betrof hier anonieme vragenlijsten die door de NZa werden opgevraagd en verwerkt. Een voorliggende vraag was of sprake was van indirect herleidbare (bijzondere) persoonsgegevens, doordat de NZa de gegevens zou kunnen koppelen aan een tweede dataset, namelijk gepseudonimiseerde declaratiegegevens waarover zij ook beschikte. De rechtbank oordeelde dat het voor de NZa in de praktijk onmogelijk was om de HoNOS+-gegevens te herleiden tot een individu. Ten eerste konden medewerkers niet tegelijkertijd bij beide datasets. Ten tweede beschikte zij niet over de sleutel om de pseudonisering van de declaratiegegevens ongedaan te maken, zelfs als zij beide datasets wel zou combineren, noch over een quantumcomputer om de versleuteling te kraken (die bestaat nog niet). Ten derde had zij geen bevoegdheid om nadere gegevens op te vragen waarmee de gepseudonimiseerde declaratiegegevens wellicht op basis van unieke combinaties aan individuen konden worden gekoppeld.

Hiernaast ging de rechtbank in op het risico dat de NZa zou worden gehackt. Daarbij overwoog zij dat *hacking* niet alleen bij wet verboden is, maar dat een eventuele hacker daarmee nog steeds slechts geanonimiseerde HoNOS+-gegevens en versleutelde declaratiegegevens in handen zou kunnen krijgen, die hij net als de NZa niet zou kunnen herleiden tot identificeerbare personen. Er bestonden volgens de rechtbank dus geen middelen waarvan redelijkerwijs valt te verwachten dat zij gebruikt worden om personen te identificeren, waardoor geen sprake was van persoonsgegevens.^b

Uit het voorgaande blijkt dat combinatie met een andere dataset een redelijk middel is om betrokkenen te identificeren, voor partijen die al toegang hebben tot de benodigde informatie, en voor partijen die over een wettige mogelijkheid beschikken om deze te verkrijgen. Dit betekent dat partijen die toegang krijgen tot de gepseudonimiseerde spelersgegevens geen toegang mogen hebben tot de (niet-gepseudonimiseerde) gegevens van individuele aanbieders (en vice versa). Voor iemand die toegang heeft tot de niet-gepseudonimiseerde gegevens van één aanbieder kunnen spelers in de gepseudonimiseerde dataset namelijk gemakkelijk identificeerbaar zijn door de twee

^a Hof Amsterdam 7 januari 2016, ECLI:NL:GHAMS:2016:146, r.o. 4.5.2-4.5.3.

^b Overigens was voor het oordeel over hacking het enkele feit dat het verboden is voor de rechtbank al genoeg [om te concluderen dat...], zie r.o. 3.12: *“Hacking is evenwel een bij wet verboden middel waarvan (dus) niet redelijkerwijs valt te verwachten dat dit door een verwerkingsverantwoordelijke of een derde wordt gebruikt om een persoon te identificeren, als bedoeld in overweging 26 van de considerans van de AVG en voornoemde jurisprudentie van het Europese Hof van Justitie.”*

datasets te koppelen, waardoor ook het speelgedrag van die spelers bij andere aanbieders kan worden gezien.^a

Identificatie op basis van de gegevens zelf

Soms is identificatie mogelijk door mensen die *zelf* over extra informatie beschikken, en daardoor iemand in de dataset kunnen herkennen, zonder dat koppeling met andere gegevens überhaupt nodig is. Denk aan situaties met familieleden of collega's die informatie hebben die anderen niet hebben, of bekende Nederlanders over wie veel persoonlijke informatie openbaar is. Dit speelde onder meer in het *Olaf*-arrest, waar in een persbericht gegevens waren opgenomen over iemands geslacht en nationaliteit, evenals het beroep van de vader en een subsidiebedrag voor een wetenschappelijk project. Het Hof oordeelde dat er in dit soort situaties '*een niet-onbeduidend risico op identificatie*' kon zijn, in casu met name door de personen die op hetzelfde wetenschappelijke gebied werkzaam waren en die de beroepsloopbaan van de betrokkene kenden.^b Om dit te voorkomen is met name de celvulling van belang.

Celvulling

Het opnemen van (te) gedetailleerde gegevens – bijvoorbeeld over leeftijd, postcode en geslacht – kan leiden tot weinig voorkomende of zelfs unieke combinaties waaruit de identiteit van een persoon kan worden afgeleid (al dan niet in combinatie met andere gegevens). Een 'cel' is een combinatie van eigenschappen en de 'celvulling' verwijst naar het aantal subjecten waarop het gegeven in de cel betrekking heeft.

Er zijn verschillende manieren om dit risico te verminderen. Deze zijn onder te delen in:

- *Afkruisen*: cellen met de lage celvulling worden leeg gelaten in de data. Wanneer ook totalen worden gegeven, moeten soms ook meerdere andere cellen worden 'afgekruid' om te voorkomen dat de waarde in de leeggelaten cel kan worden herleid.^c
- *Randomisering*: ruis toevoegen aan de gegevens, zodat deze niet altijd exact kloppen. Wanneer de ruis volledig willekeurig is, zal deze 'uitmiddelen' tot nul in de analyse, en zodoende geen invloed hebben op het onderzoeksresultaat (als er ten minste voldoende datapunten zijn). Het is uiteraard van groot belang om de wijze waarop de ruis wordt toegepast goed te ontwerpen. Wanneer

^a Dit kan zelfs wanneer niemand een sleutel bewaart om de pseudonimisering ongedaan te maken. In het speel- en stortingsgedrag van spelers zullen namelijk snel unieke combinaties zichtbaar zijn waarmee een pseudoniem uit de database aan een identificeerbare speler (in de data van een aanbieder) kan worden gekoppeld.

^b HvJ EU 7 maart 2024, C-479/22 P, ECLI:EU:C:2024:215 (*Olaf*), r.o. 60-61.

^c Zie bijvoorbeeld CBS. Kaart van 100 meter bij 100 meter met statistieken. Vlakken met minder dan 5 inwoners of minder dan 5 woningen worden niet in de dataset opgenomen.

bijvoorbeeld in een hardloop-app steeds de startlocatie van een hardloper wordt 'geruisd', zou je door het nemen van een gemiddelde over de 'geruisde' startlocaties van tientallen hardloopactiviteiten de daadwerkelijke locatie steeds preciezer kunnen achterhalen. Een oplossing hiervoor zou bijvoorbeeld kunnen zijn om één keer een geruisde locatie te bepalen voor de hardloper, en deze daarna voor alle activiteiten te gebruiken.

- *Generalisatie*: bepaalde kenmerken minder precies maken en aggregeren, zodat minder snel unieke combinatie ontstaan. Denk bijvoorbeeld aan het opnemen van een postcode-4 ("3514") in plaats van postcode 6 ("3514AX"), of het gebruik van leeftijdscategorieën ("30-35") in plaats van de exacte leeftijd ("32").^a Bij generalisatie kan rekening worden gehouden met de distributie van de onderliggende gegevens – postcodegebieden waarin minder mensen wonen zouden bijvoorbeeld kunnen worden samengevoegd om de kans op herleidbaarheid te verkleinen, terwijl in dichtbevolkte gebieden juist preciezere postcodes zouden kunnen worden gehanteerd.

Naast de risico's van onthulling van persoonsgegevens bij te lage celvulling is er het *groepsonthullingsrisico*. Hierbij is (ook bij voldoende celvulling) eenduidig een *groep* personen te koppelen aan de cel, en onthult de cel dus met hoge zekerheid een gegeven over alle individuen in de cel. Het CBS stelt als richtlijn dat een cel nooit meer dan 90% van de eenheden in de rij of kolom mag bevatten. [CBS22]

Route R2: Bijzondere persoonsgegevens hergebruiken voor onderzoek

Volgens artikel 24 UAVG is het verbod om bijzondere categorieën van persoonsgegevens te verwerken onder enkele cumulatieve voorwaarden niet van toepassing. Dit biedt ruimte om spelersgegevens te delen en te gebruiken voor onderzoek, zelfs wanneer de spelers wel identificeerbaar zijn. Aan al deze voorwaarden moet dan zijn voldaan: [UAVG art. 24]

1. De verwerking moet **noodzakelijk** zijn met het oog op wetenschappelijk of historisch onderzoek of statistische doeleinden overeenkomstig artikel 89(1) AVG.
2. Het onderzoek waarvoor de verwerking noodzakelijk is, moet **een algemeen belang** dienen.
3. Het vragen van **uitdrukkelijke toestemming moet onmogelijk zijn** of een onevenredige inspanning kosten.

^a Zie bijvoorbeeld CBS Statline: inkomen van personen, inkomensklassen, persoonskenmerken. De eigenschap 'leeftijd' is in groepen van vijf jaar gegeven.

4. Bij de uitvoering moet zijn voorzien in zodanige waarborgen dat de **persoonlijke levenssfeer van de betrokkenen niet onevenredig wordt geschaad**.

Hiernaast is ook een verwerkingsgrondslag nodig. Hieronder bespreken we eerst de voorwaarden en vervolgens de mogelijke grondslagen.

Voorwaarde 1: Noodzakelijk voor wetenschappelijk onderzoek

De eerste vraag bij het toetsen of aan deze voorwaarde wordt voldaan is of het beoogde onderzoek kwalificeert als wetenschappelijk onderzoek. In overweging 159 van de AVG wordt gesteld dat *“voor de toepassing van deze verordening de verwerking van persoonsgegevens met het oog op wetenschappelijk onderzoek ruim moet worden opgevat en bijvoorbeeld technologische ontwikkeling en demonstratie, fundamenteel onderzoek, toegepast onderzoek en uit particuliere middelen gefinancierd onderzoek omvatten.”* Hiermee is duidelijk dat het beoogde onderzoek – onafhankelijk onderzoek met wetenschappelijk of maatschappelijk doel, op het gebied van speelgedrag bij kansspelen op afstand – kwalificeert als wetenschappelijk onderzoek zoals bedoeld in deze voorwaarde. Belangrijk is wel om te beseffen dat de afbakening tot wetenschappelijk onderzoek door deze voorwaarde dus niet langer slechts een wens is, maar een randvoorwaarde.

De volgende vraag is of de verwerking van (bijzondere) persoonsgegevens *noodzakelijk* is voor dit onderzoek. Het antwoord hierop is tweeledig. In generieke zin kan worden gesteld dat het noodzakelijk is om de betreffende gegevens te verwerken om het beoogde onderzoek uit te kunnen voeren. Het doel is immers om onderzoeken mogelijk te maken die momenteel onmogelijk zijn. Ook zou dit doel niet kunnen worden bereikt door specifiek de *bijzondere* persoonsgegevens buiten de verwerking te laten, omdat die in veel gevallen juist de meest relevante gegevens zullen zijn (bijvoorbeeld voor onderzoek naar kansspelverslaving). Bij de bespreking van Route 1 in de vorige paragraaf hebben we echter beargumenteerd dat er überhaupt geen persoonsgegevens worden verwerkt als de spelers redelijkerwijs niet te identificeren zijn. Als het daadwerkelijk mogelijk is om het systeem zo in te richten dat geen sprake is van persoonsgegevens, dan kan het niet tegelijkertijd noodzakelijk zijn om (bijzondere) persoonsgegevens te verwerken.

De logische conclusie is dat de twee routes elkaar aanvullen: er moet zoveel mogelijk worden gezorgd dat de spelers niet identificeerbaar zijn (Route 1), voor zover spelers tóch identificeerbaar zijn is dat dus alleen omdat het doel anders niet kan worden bereikt. De verwerking van persoonsgegevens is dan noodzakelijk, waardoor aan deze voorwaarde voor Route 2 is voldaan. De regelgeving dwingt deze combinatie van de twee routes als het ware af. Artikel 89(1) AVG, waar deze voorwaarde naar verwijst, stelt dan ook dat de verwerking onderworpen moet zijn aan passende waarborgen voor de rechten en vrijheden van de betrokkenen. Die waarborgen houden in dat er technische en organisatorische maatregelen zijn getroffen om de inachtneming van het

beginsel van *minimale gegevensverwerking* te garanderen. Het artikel noemt pseudonimisering als voorbeeld van dit soort maatregelen.

Route 2 wordt op deze manier feitelijk een vangnet van Route 1: als (een deel van) de spelers na alle maatregelen onverhoopt toch identificeerbaar zijn, kan worden teruggevallen op artikel 24 UAVG (zolang uiteraard ook aan de overige voorwaarden wordt voldaan).

Voorwaarde 2: Het onderzoek dient een algemeen belang

Zoals eerder benoemd betreft het beoogde onderzoek onafhankelijk onderzoek met wetenschappelijk of maatschappelijk doel, op het gebied van speelgedrag bij kansspelen op afstand. Belangen daarbij zijn bijvoorbeeld het tegengaan van kansspelsverslaving of het maken van effectief beleid. Aan deze voorwaarde wordt dus voldaan als naar het doel van de onderzoeksdatabase wordt gekeken. Wel is goed om op te merken dat deze voorwaarde betekent dat dit doel dus niet kan worden gewijzigd in bijvoorbeeld het doen van commercieel onderzoek, als die wens er ooit zou zijn. Onder Route 1 zou dit wel kunnen.

Voorwaarde 3: Uitdrukkelijke toestemming vragen is onmogelijk

Zoals al eerder benoemd zou het vragen van uitdrukkelijke toestemming in de weg staan aan het inrichten van een solide onderzoeksdatabase. Toestemming vragen zou vrijwel zeker tot een *selection bias* in de data leiden (de groep subjecten die toestemming geeft is niet representatief voor de hele populatie). [Kir21] Dit betekent dat de gegevens voor veel onderzoeken niet meer goed bruikbaar zijn.^a Ook is niet goed te zeggen voor *welke* onderzoeken de gegevens dan wel of niet bruikbaar zijn, omdat niet bekend is welke factoren al dan niet een rol spelen bij de beslissing om toestemming te geven.

Vanwege het bovenstaande is het niet mogelijk om spelers uitdrukkelijk om toestemming te vragen. Voor de volledigheid merken we op dat het ook niet mogelijk is om spelers alleen toegang tot het spelaanbod te geven als zij toestemming geven om hun gegevens ook voor onderzoek te gebruiken. Toestemming dient namelijk in vrijheid te worden gegeven.^b Wanneer het gebruik van een dienst (deelname aan kansspelen) afhankelijk wordt gesteld van het geven van toestemming voor iets anders (het

^a Het is bijvoorbeeld mogelijk dat spelers die hun speelgedrag niet goed in de hand hebben minder geneigd zijn om toestemming te geven, of dat juist spelers die totaal geen problemen ervaren minder belang hechten aan onderzoek en daardoor minder snel toestemming geven. Ook zaken die niets met kansspelen te maken hebben zouden invloed kunnen hebben op het geven van toestemming, zoals vertrouwen in de wetenschap of in de overheid.

^b Zie overweging 42 Avg: *“Toestemming mag niet worden geacht vrijelijk te zijn verleend indien de betrokkene geen echte of vrije keuze heeft of zijn toestemming niet kan weigeren of intrekken zonder nadelige gevolgen.”*; zie ook [34] aant. 11 en de daar genoemde richtsnoeren van de EDPB.

gebruiken van data voor onderzoek), dan kan er niet zonder meer van worden uitgegaan dat deze in vrijheid is gegeven.

Voorwaarde 4: Geen onevenredige schade persoonlijke levenssfeer

Het gaat in deze voorwaarde om een afweging tussen het belang van de verwerking van de persoonsgegevens en de mate waarin de persoonlijke levenssfeer van de betrokkenen wordt geschaad. Hoe meer de persoonlijke levenssfeer van een of meer betrokkenen wordt geschaad, hoe groter het belang van het onderzoek moet zijn om deze inbreuk te rechtvaardigen. Of in een specifiek geval sprake is van *onevenredige* schade is in de eerste plaats ter beoordeling van de verwerkingsverantwoordelijke, die deze afweging maakt vóórdat de gegevens worden verwerkt.^a Wanneer sprake is van een wettelijke plicht is de afweging die de verwerkingsverantwoordelijke maakt evenwel beperkt tot de vraag wat noodzakelijk is om aan die plicht te voldoen. Een aanbieder kan dus niet weigeren om de wettelijk verplichte gegevens te verstrekken met het argument dat dit de betrokkenen onevenredig zou schaden. Die afweging is dan namelijk al gemaakt bij invoering van de plicht en is dus met name een politiek-bestuurlijke keuze.

De schade aan de persoonlijke levenssfeer kan en moet zo klein mogelijk worden gemaakt. Hiervoor kunnen maatregelen worden getroffen die het risico op identificatie zo klein mogelijk maken (zoals pseudonimisering) en maatregelen die de gevolgen van identificatie zo klein mogelijk maken. Voor deze beide aspecten kan onder andere worden gedacht aan het stellen van aggregatie-eisen aan publicaties, zodat geen gegevens van individuen worden gepubliceerd, en het beperken van toegang tot de onderzoeksdatabase tot geaccrediteerde onderzoeksinstituten en onderzoekers.

Grondslagen en verenigbaarheid

De AVG bepaalt dat voor verwerking van persoonsgegevens een grondslag nodig is. Dit kan een nieuwe grondslag zijn, maar het is ook mogelijk dat een nieuwe verwerking plaatsvindt die *verenigbaar* is met de doeleinden waarvoor de gegevens oorspronkelijk verzameld zijn. In dat laatste geval kan de verwerking op basis van de oorspronkelijke grondslag worden gedaan. De AVG stelt daarover dat verwerking met het oog op *“archivering in het algemeen belang, wetenschappelijk of historisch onderzoek of statistische doeleinden”* moet worden beschouwd als een met de aanvankelijke doeleinden verenigbare rechtmatige verwerking.^b Dit betekent in principe dat zowel de verdere verwerking door de aanbieders als de benodigde verwerking door een eventuele tussenpartij en door de onderzoekers kunnen plaatsvinden op grond van de oorspronkelijke grondslag (de wettelijke plicht op grond waarvan de aanbieders de gegevens hebben verzameld). Al deze verwerkingen worden immers gedaan met het oog

^a Als een betrokkene later vindt dat wel sprake is van een onevenredige schade en partijen er niet uitkomen, wordt de beoordeling gedaan door de Autoriteit Persoonsgegevens en/of uiteindelijk de rechter.

^b Overweging 50 AVG.

op wetenschappelijk onderzoek. De benodigde verwerkingen kunnen echter ook op grond van nieuwe grondslagen plaatsvinden.

In het geval van een nieuwe grondslag zijn er, afhankelijk van de partij en gekozen oplossingsrichting, drie grondslagen die voor de hand liggen:

- de verwerking is noodzakelijk om te voldoen aan een wettelijke plicht (art. 6(1)(c) AVG),
- de verwerking is noodzakelijk voor de vervulling van een taak van algemeen belang (art. 6(1)(e) AVG),
- de verwerking is noodzakelijk voor de behartiging van een gerechtvaardigd belang (art. 6(1)(f) AVG).

Indien medewerking van partijen wettelijk wordt geborgd, is het aannemelijk dat de verstrekking van de gegevens door de aanbieders zal plaatsvinden op grond van een specifieke wettelijke plicht. Dat kan artikel 31m lid 5 Wok zijn, of een nieuwe bepaling in de Wok, of de Wet CBS.^a Onderzoekers die de gegevens gaan gebruiken kunnen hun verwerking baseren op de grondslag van een gerechtvaardigd belang. Voor een eventuele intermediair is mogelijk dat deze belast wordt met een taak van algemeen belang, waardoor die grondslag van toepassing is, of op grond van een gerechtvaardigd belang (als het geen overheidsorganisatie is).

3.3 Borgen van medewerking

3.3.1 Juridische obstakels

In de vorige paragraaf is besproken onder welke voorwaarden de spelersgegevens gedeeld *mogen* worden. Daarnaast is echter ook nodig dat aanbieders dit daadwerkelijk doen. Zoals eerder besproken biedt artikel 31m lid 5 Wok een verplichting voor aanbieders om gegevens te verstrekken, maar deze verplichting is niet breed genoeg om in de volledige onderzoeksbehoefte te voorzien. Het artikel bevat een verplichting voor aanbieders om *geanonimiseerde* spelersgegevens ter beschikking te stellen, ten behoeve van onderzoek naar *kansspelverslaving*. Aanbieders zijn dus niet verplicht om gegevens te verstrekken die niet-geanonimiseerd zijn, zelfs als ze dit op grond van artikel 24 UAVG wel zouden mogen. Ook zijn ze niet verplicht om (al dan niet geanonimiseerde) gegevens te verstrekken voor onderzoek dat niet over kansspelverslaving gaat. Daarnaast is de verplichting momenteel op een specifieke manier uitgewerkt, die onder meer inhoudt dat onderzoekers de gegevens één-op-één bij aanbieders opvragen (zie paragraaf 2.3).

^a Het gebruik van een bevoegdheid die het CBS heeft om gegevens op te vragen komt in paragraaf 3.3.2 aan de orde.

Voor nieuwe oplossingsrichtingen die in dit onderzoek worden besproken, waarmee koppelbare gegevens beschikbaar kunnen worden gesteld, kan dus niet altijd worden teruggevallen op de huidige (uitwerking van de) verplichting op basis van artikel 31m lid 5 Wok. Ook kan er niet zonder meer van worden uitgegaan dat kansspelaanbieders vrijwillig aan een oplossing meewerken, aangezien zij (1) vermoedelijk kosten zullen moeten maken om hun gegevens beschikbaar te stellen en (2) meer maatschappelijk en wetenschappelijk onderzoek naar kansspelen niet vanzelfsprekend in hun belang is. Het ligt daarom in de rede dat er iets moet worden gedaan om aanbieders te motiveren, dan wel te verplichten, om aan de gekozen oplossing mee te werken.

3.3.2 Oplossingen voor de juridische obstakels

Om de aanbieders te bewegen de voorziene data te delen, zien we drie routes, die we hieronder nader uitwerken. Deze routes zijn genummerd en voorzien van voorvoegsel 'M' om ze te onderscheiden van de routes met betrekking tot rechtmatigheid, die we verderop bespreken. Achtereenvolgens gaan we in op *zelfregulering* (route M1), *overheidsregulering* (M2) en *dataverzameling door het CBS* (M3).

Route M1: Zelfregulering

Er is niet altijd overheidsregulering nodig om een sector te reguleren. Zelfregulering, waarbij de sector zichzelf committeert aan bepaalde regels, kan onder omstandigheden zelfs de voorkeur hebben. In de literatuur worden onder meer als (mogelijke) voordelen genoemd dat zelfregulering flexibeler is dan wetgeving en kan leiden tot een grotere bereidheid tot naleving. [van09] De reguleringskosten en nalevingskosten zouden bij zelfregulering daarnaast lager zijn dan bij wetgeving. [Baa04]

Zelfregulering is echter niet altijd passend. Hoewel zelfregulering tot op zekere hoogte kan worden 'afgedwongen' door een sector wetgeving boven het hoofd te hangen, is dit geen garantie dat (een poging tot) zelfregulering ook zal slagen. Het kan bijvoorbeeld falen doordat te weinig marktpartijen zich aan de regels committeren, of doordat handhaving van de opgestelde zelfregulering moeilijker is dan de handhaving van wetgeving zou zijn. In eerder onderzoek zijn factoren geïdentificeerd [van09], en later getoetst [van11], die verklarend zijn geweest voor het slagen van zelfregulering op verschillende terreinen en die voorzichtig ook ex ante gebruikt kunnen worden om te beoordelen of een casus geschikt lijkt voor zelfregulering. Deze factoren zijn vastgelegd in een beoordelingskader.^a We bespreken achtereenvolgens vier factoren uit dit kader: (1) de *organisatiegraad van de sector*, (2) *aard van het te dienen belang*, (3) *effecten op omzet en winst*, en (4) *draagvlak onder de kansspelaanbieders*.

^a Zie ook het [Kenniscentrum voor beleid en regelgeving](#) van de Rijksoverheid, waar naar dit kader wordt verwezen.

Organisatiegraad van de sector

Een belangrijke factor die in deze onderzoeken naar voren is gekomen is de *organisatiegraad* van de sector. De organisatiegraad in de kansspelsector is vrij hoog: het gaat om enkele tientallen aanbieders, die vanwege het vergunningenstelsel en bestaande regels gemakkelijk te inventariseren en te bereiken zijn. Ook zijn veel aanbieders aangesloten bij een branchevereniging. Deze hoge organisatiegraad vergroot de kans van slagen van zelfregulering. Dit lijkt echter de enige factor uit het kader te zijn die in het geval van spelersdata vóór zelfregulering pleit.

Aard van het door zelfregulering te dienen belang (privaat versus publiek of politiek)

Een tweede aspect uit het selectiekader uit [van09] is de aard van het te dienen belang. Hierover wordt het volgende opgemerkt:

Als het object van zelfregulering eenduidig (mede) het privaat belang van betrokkenen betreft (bijv. keurmerkfunctie, uitstralen kwaliteit) is de kans op totstandkoming en handhaving van zelfregulering groter dan wanneer het (uitsluitend) een publiek c.q. politiek belang betreft dat niet convergeert met het privaat belang. [van09 p. 12]

Het belang dat wordt gediend met het beschikbaar maken van spelersgegevens voor onderzoek is vrijwel uitsluitend een publiek belang. Het onderzoek waarvoor de gegevens beschikbaar moeten komen is immers afgebakend tot onafhankelijk wetenschappelijk of maatschappelijk onderzoek. De belangen die hiermee worden gediend zijn onder meer betere consumentenbescherming en verslavingspreventie, betere monitoring en effectiever beleid. Dit zijn bij uitstek publieke belangen die niet convergeren met het private belang.

Effecten op omzet en winst

Een derde aspect uit het selectiekader uit [van09] is het effect op omzet en winst. Hierover wordt het volgende opgemerkt:

Als de branche geen belang ziet in het meewerken aan het beleidsdoel én de branche verwacht dat inpassing in bedrijfsvoering winstdrukkend zal zijn, is ontstaan van zelfregulering c.q. de kans van slagen gering. [van09 p. 13]

In de periode oktober 2023 tot september 2024 was 1% van de spelersaccounts verantwoordelijk voor 43% van de *gross gaming revenue (GGR)* van online kansspel-aanbieders, dat wil zeggen: 43% van al het geld dat spelers hebben verloren kwam van 1% van de spelersaccounts. [Kan251] De inkomsten van de aanbieders zijn dus sterk afhankelijk van een zeer klein deel van de spelers dat veel geld verliest (€ 2500+ per maand per account). Hoewel een deel van hen zich deze verliezen zal kunnen veroorloven, zal een deel zijn speelgedrag niet in de hand hebben en hierdoor in de problemen raken. Dit zijn de gevallen die met effectievere consumentenbescherming en

verslavingspreventie minder vaak voor zullen komen. Het bereiken van deze doelen zal dus winstdrukkend zijn voor de bedrijfsvoering van online kansspelaanbieders. Daarnaast zal ook het beschikbaar maken van de gegevens zelf vermoedelijk al gepaard gaan met invoeringskosten voor de aanbieders.

Draagvlak onder kansspelaanbieders

Een vierde aspect uit het selectiekader uit [van09] is het draagvlak onder (in deze casus) kansspelaanbieders. Hierover wordt het volgende opgemerkt:

Hoe groter het draagvlak binnen de branche voor de inhoud en vorm van de normering, des te groter de kans op totstandkoming en naleving van zelfregulering. Negatief geformuleerd: bij freerider probleem c.q. cowboy-praktijken zal die kans kleiner zijn. Volledige naleving is ook met groot draagvlak niet gegarandeerd. [van09 p. 13]

We hebben in dit onderzoek slechts één aanbieder en één brancheorganisatie gesproken en hebben het draagvlak binnen de branche verder niet onderzocht. Op basis van eerder onderzoek kunnen we echter het volgende zeggen over het draagvlak voor maatregelen die goed zijn voor de maatschappij, maar minder goed voor het bedrijfsresultaat van aanbieders. Hoewel er aanbieders zullen zijn die het maatschappelijk belang van onderzoek en het moreel belang van (o.a.) verslavingspreventie zwaarder vinden wegen dan hun eigen financiële belang, en zich dus zullen willen committeren aan zelfregulering, is het aannemelijk dat er ook aanbieders zijn voor wie dit niet geldt. In eerder onderzoek is op meerdere onderwerpen van regulering in de kansspelsector gebleken dat er zowel aanbieders zijn die zich netjes aan de regels houden als aanbieders die het grijze gebied opzoeken of regels overtreden ten behoeve van hun eigen bedrijfsbelang.^a

Wanneer niet alle aanbieders zich aan de zelfregulering committeren en meewerken aan het beschikbaar stellen van spelersdata voor onderzoek, ontstaat er een groot risico: aanbieders die niet meedoen kunnen gaan adverteren met het feit dat spelersdata bij hen niet wordt gedeeld met onderzoekers. Dit kan bepaalde spelers overhalen om juist bij die aanbieders te gaan spelen, wat afbreuk zou doen aan de representativiteit van de gegevens die wél beschikbaar zijn voor onderzoek en daarmee aan de betrouwbaarheid en bruikbaarheid van de onderzoeksresultaten. Indien andere aanbieders vervolgens zouden menen dat ze hun eigen concurrentiepositie verzwakken door mee te werken aan gegevensdeling voor onderzoek, zou dit ertoe kunnen

^a Uit de evaluatie van de Wet Koa blijkt bijvoorbeeld dat dit geldt voor de uitvoering van de zorgplicht en de regels omtrent werving en reclame, zie [36]. Ook in onderzoek naar het gebruik van promotionele kansspelen kwam dit verschil tussen aanbieders sterk naar voren. [37]

leiden dat ook die aanbieders besluiten niet langer mee te werken aan de zelfregulering.

Op basis van het voorgaande lijkt zelfregulering niet de aangewezen route om spelersgegevens beschikbaar te maken voor onderzoek. Vanwege de conflicterende belangen is er een risico dat de zelfregulering moeilijk tot stand komt, zonder volledige dekking in de sector, en slecht wordt nageleefd. Niet-volledige dekking zou bovendien kunnen leiden tot minder betrouwbare onderzoeksresultaten. Wellicht ten overvloede merken we op dat het uiteraard mogelijk is om een poging tot zelfregulering te doen en, als deze faalt, alsnog over te gaan tot regulering. Gelet op de huidige zorgen in de kansspelsector en de grote behoefte aan op data gebaseerd beleid, is het echter de vraag of men dit risico op enkele jaren vertraging moet willen nemen.

Route M2. Aanpassing kansspelregelgeving

Tegen de achtergrond van de naar verwachting beperkte effectiviteit van zelfregulering, lijkt overheidsregulering voor de hand te liggen om de medewerking van kansspelaanbieders te waarborgen. We hebben het dan concreet over het creëren van een (wettelijke) *verplichting* voor kansspelaanbieders om de betreffende spelersgegevens beschikbaar te maken voor onderzoek, in overeenstemming met de (nog te kiezen) oplossing. Gelet op het huidige wettelijke kader van kansspelen, waarin de Wok de primaire wet is, is het logisch om dit in de Wok of onderliggende regelgeving te regelen.

Als de verplichting in de Wok zelf wordt opgenomen, door artikel 31m lid 5 te wijzigen of een nieuwe bepaling op te nemen, kan de wetgever daar uiteraard in opnemen wat hij wil. Er kan dus desgewenst worden geregeld dat (ook) spelersgegevens gedeeld moeten worden die niet kunnen worden beschouwd als geanonimiseerde gegevens, of dat gegevens ook gedeeld moeten worden voor een ander doel dan onderzoek naar kansspelverslaving. Vanwege de kortere doorlooptijd heeft het echter de voorkeur om enkel lagere regelgeving te hoeven wijzigen. Dit laatste heeft echter de beperking dat moet worden aangesloten op een bestaande wettelijke bepaling die de bevoegdheid creëert om de gewenste verplichting in lagere regelgeving te regelen.^a Het enige aanknopingspunt dat we hiervoor zien is het bestaande artikel 31m leden 5 en 6 Wok. De verplichting van lid 5 zou in het onderliggende besluit anders uitgewerkt kunnen worden dan momenteel het geval is, maar men blijft dan wel gebonden aan de bewoording van het huidige lid. De mogelijkheden binnen die route bespreken we hier.

Het vijfde lid van het artikel verplicht kansspelaanbieders om bepaalde gegevens geanonimiseerd beschikbaar te stellen voor onderzoek naar kansspelverslaving:

^a De wetgever delegeert deze bevoegdheid dan aan de regering.

De vergunninghouder stelt gegevens en analyses als bedoeld in het eerste lid geanonimiseerd beschikbaar voor onderzoek naar kansspelverslaving.

Het zesde lid regelt vervolgens dat deze verplichting nader kan worden uitgewerkt in lagere regelgeving. In dit geval is dat in het Besluit werving, reclame en verslavingspreventie kansspelen (Besluit Wrvk [Wrvk]) en een onderliggende specifieke regeling.^a Daarin is momenteel de procedure vastgelegd waarmee onderzoekers gegevens kunnen opvragen bij individuele aanbieders, maar dit kan worden gewijzigd in, of aangevuld met, een procedure die beter tegemoetkomt aan de onderzoeksbehoefte. In plaats van (alleen) te regelen dat onderzoekers een direct verzoek aan individuele aanbieders kunnen doen, kan bijvoorbeeld (ook) worden geregeld dat aanbieders een bepaalde set gegevens maandelijks uploaden in een centrale onderzoeksdatabase.

Zoals zojuist benoemd moet een eventuele nieuwe uitwerking wel aansluiten op het huidige artikel 31m lid 5 Wok. Het is immers een uitwerking van die bepaling. Dat betekent dat de gegevens *geanonimiseerd* moeten zijn en dat zij gedeeld worden *voor onderzoek naar kansspelverslaving*. Op het eerste gezicht lijken dit belangrijke beperkingen met het oog op de onderzoeksbehoefte. Men wil immers dat de spelersgegevens van verschillende aanbieders koppelbaar zijn en de onderzoeksbehoefte is breder dan enkel het thema kansspel *verslaving*. Juridisch lijkt er echter ruimte te zijn om toch in deze behoefte te voorzien, zonder de bewoordingen van lid 5 te wijzigen. Dat de gegevens ook in een potentiële nieuwe uitwerking van de verplichting geanonimiseerd moeten zijn, betekent namelijk niet per definitie dat zij niet koppelbaar over aanbieders zijn. ‘Geanonimiseerd’ wil zeggen dat niet langer sprake is van persoonsgegevens, en zoals besproken in de paragraaf over rechtmatigheid (3.2) betekent dit dat het redelijkerwijs niet mogelijk en te verwachten moet zijn dat de identiteit van de desbetreffende natuurlijke persoon wordt achterhaald. Zoals in die paragraaf uiteengezet kan dit ook worden bereikt wanneer gebruik wordt gemaakt van pseudoniemen waarmee spelersgegevens koppelbaar zijn over aanbieders, zolang er passende maatregelen worden genomen waardoor de identiteit van spelers naar verwachting niet redelijkerwijs kan worden achterhaald.

Ook de afbakening tot onderzoek naar kansspelverslaving is wellicht geen onoverkomelijke beperking. Op grond van artikel 31m lid 5 Wok zijn kansspelaanbieders enkel verplicht om gegevens aan te leveren als het gaat om onderzoek naar kansspelverslaving, maar dat betekent niet direct dat die gegevens daarna niet óók voor ander onderzoek gebruikt mogen worden. Vanuit de AVG bezien kan dit bijvoorbeeld onder doelverenigbaarheid en de Wok zelf staat er niet aan in de weg. Er lijkt dus ruimte te zijn om de verplichting zo uit te werken dat aanbieders de gegevens moeten aanleveren aan een onderzoeksdatabase die primair bedoeld is voor onderzoek naar

^a in art. 17 Besluit Wrvk en in de Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving. [6]

kansspelsverslaving, maar die daarnaast ook gebruikt kan worden voor ander kansspelonderzoek.^{a,b}

Impact van een restrictieve interpretatie van 31m

Als de restrictieve wetsinterpretatie wordt aangehouden (zie paragraaf 3.2 onder de bespreking van 31m lid 5 Wok), waarin de wetgever heeft willen beperken wat er met de spelersgegevens gebeurt, biedt het wijzigen van de lagere regelgeving minder ruimte met betrekking tot het onderzoeksdoel van kansspelsverslaving. Als de wetgever streng heeft willen beperken wat er met de gegevens mag gebeuren, dan kunnen de gegevens die worden gedeeld ten behoeve van onderzoek naar kansspelsverslaving daarna niet ook gebruikt worden voor onderzoek naar andere onderwerpen. In dat geval zou een wijziging van de Wok zelf nodig zijn om de gewenste gegevensdeling voor andere onderzoeksdoeleinden mogelijk te maken en medewerking van aanbieders te borgen.

Route M3: Verzameling door het CBS

Er bestaat via het CBS een andere route om de gewenste gegevens beschikbaar te maken voor wetenschappelijk onderzoek, waarbij medewerking van kansspelaanbieders geborgd is met een (bestaande) wettelijke plicht. CBS kan bij kansspelaanbieders gegevens opvragen ten behoeve van de uitoefening van de hem bij wet opgedragen taken (daaronder begrepen opdrachten van de minister). (Wet CBS, art. 33 lid 3). Het CBS kan de gegevens vervolgens beschikbaar maken voor onderzoekers. Hiervoor is geen wetswijziging nodig, maar wel een AMvB. We lichten dit hieronder toe.

Het CBS verzamelt en beheert een grote hoeveelheid data, die zij gebruikt voor het bepalen en publiceren van statistiek, en daarnaast beschikbaar maakt voor onderzoekers. Om zijn wettelijke taak uit te voeren heeft het CBS de bevoegdheid om gegevens op te vragen bij, onder meer, bepaalde categorieën van ondernemingen, die verplicht zijn om deze binnen een vastgestelde termijn aan te leveren. (Wet CBS, art. 33). Bij AMvB kan worden geregeld dat het CBS ook bij kansspelaanbieders gegevens op kan vragen. In de AMvB wordt dan vastgelegd welke gegevens kunnen worden opgevraagd en binnen welke termijn die geleverd moeten worden (Wet CBS, art. 33). Het feit dat het om bijzondere persoonsgegevens gaat staat hieraan niet in de weg. Het CBS kan ten

^a Merk op dat de gegevens geanonimiseerd moeten zijn, waardoor de AVG geen beperkingen stelt.

^b Uiteraard kan ook worden gediscussieerd over de afbakening van het begrip 'kansspelsverslaving' in deze context – onderzoek naar kansspelgedrag is bijvoorbeeld al snel relevant voor kansspelsverslaving.

behoefte van statistische doeleinden bijzondere categorieën van persoonsgegevens verwerken (Wet CBS, art. 35).^a

Tijdens de coronapandemie is deze route voorgesteld om vaccinatiegegevens van het RIVM beschikbaar te maken voor statistisch onderzoek naar oversterfte door het CBS en (vervolgens) voor wetenschappelijk onderzoek door academici.^b De AP was destijds terughoudend in haar advies over deze route, maar stelde dat het in haar ogen toegestaan was zolang alle gegevens daadwerkelijk werden opgevraagd met als doel 'statistisch onderzoek door het CBS zelf':

"De AP merkt hierbij op dat het vanzelfsprekend niet zo kan zijn dat het CBS op basis van artikel 33 Wet CBS bij (overheids)instanties persoonsgegevens uitvraagt ten behoeve van wetenschappelijk onderzoek niet zijnde statistisch onderzoek door het CBS zelf. Om in ieder geval de indruk van détournement de pouvoir te voorkomen, is bij gebruik van deze bevoegdheid door het CBS oplettendheid geboden. Immers de uitvraag door het CBS van gegevens mag niet gericht zijn op verkrijgen van gegevens voor wetenschappelijk onderzoek maar moet gericht zijn op gegevens voor statistisch onderzoek door het CBS zelf, dat resulteert in openbaar te maken statistieken in lijn met de wettelijke taak." [APCovid]

Uit een latere brief van de Minister van Volksgezondheid, Welzijn en Sport blijkt dat men ervoor heeft gekozen om deze route daadwerkelijk in te zetten. Momenteel zijn de gegevens dan ook opgenomen in de CBS Microdata catalogus.^c Onderstaand citaat vat de werkwijze samen, die met enige aanpassingen in theorie ook mogelijk zou moeten zijn voor spelersgegevens van kansspel aanbieder.

"Onderdeel van deze werkwijze is dat het RIVM de data via een beveiligde verbinding bij het CBS aanlevert en dat de data na ontvangst direct gepseudonimiseerd worden binnen de beveiligde CBS-omgeving. Het CBS gebruikt deze data uitsluitend voor het uitvoeren van statistisch onderzoek. Het CBS stelt de data in gepseudonimiseerde vorm ook aan onderzoekers van door het CBS gemachtigde instellingen beschikbaar in de remote-accessomgeving ten behoeve van wetenschappelijk en/of statistisch onderzoek, overeenkomstig artikel 41 en artikel 42 van de Wet op het CBS. Om gebruik te kunnen maken van de data zullen de wetenschappers onder meer een onderzoeksvoorstel moeten indienen bij ZonMW. Ook

^a Voor de kansspel aanbieder zou gelden dat ze de bijzondere persoonsgegevens mogen delen op grond van art. 24 uAVG (zie paragraaf 3.2.2), in combinatie met de grondslag van een wettelijke plicht op basis van de Wet CBS.

^b Zie [29]. Het ging hier weliswaar om het opvragen van gegevens bij een overheidsinstantie (RIVM) in plaats van bij ondernemingen (kansspel aanbieder), maar de het CBS heeft beide bevoegdheden op basis van artikel 33 Wet CBS [32]).

^c Onder de naam *Documentatie COVID-19 vaccinatiegegevens uit het COVID-vaccinatie Informatie- en Monitoringssysteem (CIMS)*. [28]

zullen de publicaties van de onderzoeken moeten voldoen aan de regels van het CBS, waaronder een controle ter voorkoming van onthulling van gegevens van individuen. Hiermee komen de vaccinatiedata van het RIVM beschikbaar voor het onafhankelijk wetenschappelijk onderzoek [...].” [Min22 p. 7, par 3]

De minister ging in zijn brief niet expliciet in op het advies van de AP en het onderscheid tussen de doeleinden ‘statistisch onderzoek door het CBS’ en ‘wetenschappelijk onderzoek’, maar stelde slechts dat de gegevens voor beide doelen beschikbaar werden gemaakt. Het doen van statistisch onderzoek naar kansspelen, ten behoeve van beleid en wetenschap, lijkt echter probleemloos onder te brengen in de wettelijke taken van het CBS. Artikel 3 van de Wet CBS luidt namelijk als volgt:

1. *Het CBS heeft tot taak het van overheidswege verrichten van statistisch onderzoek ten behoeve van praktijk, beleid en wetenschap en het openbaar maken van de op grond van zodanig onderzoek samengestelde statistieken.*
2. *Het CBS bevordert:*
 - a. *een statistische informatievoorziening van overheidswege die voorziet in de behoeften van praktijk, beleid en wetenschap;*

[...]

Ervan uitgaande dat het CBS in het kader van zijn wettelijke taken ook statistisch onderzoek kan doen naar de kansspelmarkt, kan het CBS de gegevens dus opvragen en vervolgens ook ter beschikking stellen aan onderzoeksinstellingen.

Zoals eerder benoemd kan het CBS de gegevens opvragen bij de kansspelaanbieders zelf als de daarvoor benodigde AMvB is afgegeven. Een alternatieve optie, ook mogelijk gemaakt door de Wet CBS, is om de gegevens op te vragen bij de Ksa. Dan zou geen AMvB nodig zijn.^a Er zijn echter beperkingen die deze optie minder aantrekkelijk maken. Ten eerste zijn in dat geval enkel de gegevens op te vragen waarover de Ksa zelf beschikt, zijnde de CDB-data, terwijl bij de aanbieders meer opgevraagd zou kunnen worden (zoals bijvoorbeeld ruwe data uit systemen – die relevant zou kunnen zijn voor onderzoek waarbij ook ‘in’ de spellen wordt gekeken). Ten tweede zijn spelers in de CDB-data gepseudonimiseerd met door de aanbieders gekozen pseudoniemen, waardoor de gegevens niet koppelbaar zijn. Er zou dus alsnog iets moeten wijzigen in de wijze waarop aanbieders de gegevens pseudonimiseren, bijvoorbeeld door aanbieders gebruik te laten maken van pseudoniemen die door de Ksa zijn opgesteld. Ten derde wordt het lastig of onmogelijk om het datamodel te wijzigen ten behoeve van onderzoek, aangezien gebruik wordt gemaakt van gegevens die worden verzameld ten

^a De Ksa is een ZBO op het niveau van de centrale overheid, dus CBS kan de gegevens opvragen o.g.v. art. 33 lid 1 onderdeel c Wet CBS. [32]

behoefte van toezicht. Het ligt daarom meer voor de hand om het CBS de gegevens op te laten vragen bij de aanbieders zelf, ondanks dat daarvoor een AMvB vereist is.

3.4 Taakstelling Ksa en doelstelling CDB

3.4.1 Juridische obstakels

De taakstelling van de Ksa en de doelstelling van de controledatabank (CDB) vormen juridische obstakels, gelet op de eerder in paragraaf 2.3 benoemde ontwerpdoelen.

Taakstelling Ksa beperkt mogelijkheden voor eigen onderzoek

De Wet op de Kansspelen stelt dat er een kansspelautoriteit is, en kent daar zeven wettelijke taken aan toe (Wok, art. 33b, opmaak gewijzigd door Dialogic):

[De Raad van Bestuur van de kansspelautoriteit] heeft, tenzij bij of krachtens deze wet anders is bepaald, tot taak:

- *het verlenen, wijzigen, schorsen en intrekken van vergunningen voor de diverse vormen van kansspelen, exploitatievergunningen en modeltoelatingen voor speelautomaten,*
- *het beheer van het register, bedoeld in artikel 33h, [CRUKS, red.]*
- *het bevorderen van het voorkomen en het beperken van kansspelverslaving*
- *het geven van voorlichting en informatie*
- *het bestrijden van niet toegestaan kansspelaanbod*
- *het toezicht op de naleving van de toepasselijke wet- en regelgeving en de vergunningen, de handhaving daarvan*
- *alsmede het tegengaan en beperken van manipulatie met sportwedstrijden [...].*

Het uitvoeren van maatschappelijk of wetenschappelijk onderzoek is geen taak van de Ksa. Het is de Ksa echter niet verboden onderzoek uit te voeren ten behoeve van de uitvoering van haar wettelijke taken. De Ksa voert in de praktijk ook onderzoek uit, in het kader van haar toezichtstaak. [Kan251] De wettelijke taakstelling maakt dat de Ksa echter geen onderzoek kan doen dat niet direct in relatie staat tot de taken (bijvoorbeeld meer vooruitkijkend, of onderzoek dat is gericht op beleidsvorming).

Besluit Koa: CDB is niet bedoeld voor onderzoek

In het Besluit Koa [Min211] wordt gedefinieerd dat vergunninghouders een controledatabank moeten bijhouden, wat daar vervolgens in moet staan, en welke maatregelen voor informatiebeveiliging moeten worden genomen.^a De controledatabank is ingericht voor toezichtsdoeleinden en enkel toezichthouders hebben er toegang toe. In de

^a Tevens relevant in het kader van dit onderzoek is ook dat dit artikel toestaat dat de categorieën van informatie in de controledatabank met een regeling kunnen worden vastgesteld (artikel 5.3, lid 7)

toelichting bij artikel 5.3 Besluit Koa wordt zelfs expliciet benoemd dat de CDB niet is bedoeld om wetenschappelijk onderzoek mee uit te voeren:

De controledatabank is niet bedoeld als middel voor (grootschalig) wetenschappelijk onderzoek naar bijvoorbeeld kansspelverslaving.

Zelfs als het takenpakket van de Ksa wordt aangepast, kan de controledatabank dus niet zomaar beschikbaar worden gesteld voor onderzoek. Daarvoor zou ook de regelgeving over de CDB zelf moeten worden aangepast. Niet uitgesloten is echter dat de Ksa separaat een gegevensdeling faciliteert op basis van het CDB-*datamodel*, waarbij dezelfde datasets (in ieder geval *de jure*) opnieuw worden aangeleverd door de aanbieders ten behoeve van onderzoek. De CDB zelf blijft dan slechts toegankelijk voor toezicht, maar er wordt dan nieuwe databank gecreëerd ten behoeve van onderzoek, met (in eerste instantie) dezelfde gegevens.

3.4.2 Oplossingen voor de juridische obstakels

Route T1. Wijzigen taakstelling Ksa en besluit Koa

Om de mogelijkheid tot het uitvoeren van onderzoek voor de Ksa te verbreden, zou de wettelijke taakstelling van de Ksa (Wok art. 33b) kunnen worden aangepast.^a Het wijzigen van de taakstelling is eveneens nodig wanneer de Ksa zou moeten optreden als intermediair bij het delen van data, en/of als partij die de pseudoniemen voor onderzoek faciliteert, en/of andere vormen van datadeling zou moeten gaan faciliteren.

Voor wat betreft de toegang tot, en doelstelling van, de controledatabank, ligt het ons inziens niet voor de hand dit te wijzigen ten behoeve van onderzoek. De regelgeving over de CDB staat namelijk niet in de weg aan het separaat delen van soortgelijke gegevens ten behoeve van onderzoek. Het ligt meer voor de hand om de datadeling voor onderzoek separaat van de datadeling ten behoeve van toezicht uit te werken. Op die manier blijven onderzoek en toezicht gescheiden en kan een eventuele onderzoeks-databank worden aangepast zonder de CDB aan te passen en vice versa. Daarbij kan echter wel nauw worden aangesloten bij de bestaande werkwijzen en datamodellen.

3.5 Conclusie

In dit hoofdstuk worden de drie belangrijkste juridische obstakels en oplossingen daarvoor bij de voorgenomen datadeling uiteengezet. Het eerste obstakel is

^a Bij het formuleren van een aangepaste taakstelling kan wellicht worden gekeken naar de wijze waarop dit is gedaan voor de Autoriteit Consument en Markt, die een separate taakstelling heeft betreffende onderzoek (zie Instellingswet ACM, artikel 2, lid 4).

rechtmatigheid van de datadeling, het tweede is **medewerking** aan de datadeling, en het derde is **de taakstelling van de Ksa**.

Taakstelling van de Ksa

Over de taakstelling van de Ksa kunnen we kort zijn: om mogelijk te maken dat de Ksa onderzoek doet dat verder gaat dan 'onderzoek ten behoeve van toezicht', moet de taakstelling van de Ksa in artikel 33b Wok worden aangepast (Route T1). Over de twee andere thema's trekken we hieronder onze conclusies.

Rechtmatigheid datadeling

Met betrekking tot het rechtmatig delen van spelersgegevens voor onderzoek hebben we twee mogelijke juridische obstakels geanalyseerd. Ten eerste speelde de vraag of de kansspelwetgeving in de weg staat aan het delen van spelersgegevens op andere wijze dan op grond van artikel 31m lid 5 Wok, dat aanbieders verplicht om gegevens (geanonimiseerd) beschikbaar te stellen ten behoeve van onderzoek naar kansspelverslaving. Ten tweede is er het verbod uit de AVG op het verwerken van bijzondere persoonsgegevens.

De kansspelwetgeving staat volgens ons niet in de weg aan het delen van niet-geanonimiseerde spelersgegevens. We kunnen dit niet met zekerheid stellen, omdat de regelgeving op dit punt niet geheel eenduidig is. De kansspelregelgeving bevat (art. 31m Wok) een *verplichting voor aanbieders om mee te werken aan* een verzoek om geanonimiseerde gegevens ten behoeve van onderzoek naar kansspelverslaving, maar verbiedt aanbieders volgens ons niet om niet-geanonimiseerde gegevens te verstrekken, noch om gegevens te verstrekken voor onderzoek naar een ander thema dan kansspelverslaving.

Het gevolg is dat alleen nog het verbod uit de AVG op het verwerken van bijzondere persoonsgegevens resteert als juridisch obstakel voor het rechtmatig beschikbaar stellen van de gegevens. Er zijn twee manieren om met het verbod om te gaan:

- **Route R1: anonimiseren.** Het is mogelijk om spelersgegevens te delen die wél koppelbaar zijn tussen aanbieders, maar níet meer (redelijkerwijs) te herleiden zijn tot identificeerbare personen, en dus kwalificeren als geanonimiseerde gegevens. Dit kan worden bereikt met een combinatie van technische en juridische maatregelen. De benodigde technische maatregelen om heridentificatie te voorkomen, zoals generaliseren en randomiseren, zullen zeer waarschijnlijk leiden tot een lager detailniveau en lagere kwaliteit van de gegevens, vergeleken met een situatie waarin wel persoonsgegevens gedeeld mogen worden.
- **Route R2: artikel 24 UAVG.** Om wetenschappelijk onderzoek mogelijk te maken dat niet uitgevoerd kan worden met geanonimiseerde gegevens, kan gebruik worden gemaakt van artikel 24 UAVG. Dit artikel biedt een uitzondering op het verbod om bijzondere persoonsgegevens te verwerken, voor wanneer dit

noodzakelijk is voor wetenschappelijk of statistisch onderzoek met een algemeen belang. Ook in deze route moeten technische en organisatorische maatregelen worden genomen, namelijk om minimale gegevensverwerking te bewerkstelligen.

Of Route R1 of R2 wordt gebruikt is feitelijk geen keuze, maar is afhankelijk van de te kiezen oplossingsrichting en de daarbij gestelde doelen. Als het binnen de gekozen oplossingsrichting redelijkerwijs *mogelijk* is om de doelstellingen te bereiken met geanonimiseerde gegevens, dan *moeten* de gegevens worden geanonimiseerd. Daarmee is automatisch Route R1 van toepassing. Is het evenwel niet mogelijk om de doelstellingen te bereiken met geanonimiseerde gegevens, dan kan gebruik worden gemaakt van Route R2.

Restrictieve interpretatie van artikel 31m Wok

Zoals benoemd is niet geheel eenduidig of de Wok in de weg staat aan het delen van spelersgegevens op andere wijze dan de bestaande verplichting van artikel 31m lid 5 voorschrijft. Onze interpretatie is dat de regelgeving hieraan niet in de weg staat, maar een restrictieve interpretatie is ook verdedigbaar. Deze komt erop neer dat spelersgegevens *enkel* mogen worden gedeeld wanneer de verstrekking voldoet aan artikel 31m lid 5, met de volgende gevolgen:

- Het is niet toegestaan om niet-geanonimiseerde gegevens te delen, zelfs als de AVG hier niet aan in de weg zou staan. Route R2 is daarmee geen optie meer.
- Het is niet toegestaan om geanonimiseerde gegevens te delen voor andere doeleinden dan onderzoek naar kansspelsverslaving. Route R1 is daarmee beperkt mogelijk.

Als deze interpretatie wordt aangehouden is het nog steeds mogelijk om koppelbare spelersgegevens beschikbaar te maken, zoals toegelicht kunnen geanonimiseerde gegevens immers nog steeds koppelbaar zijn, maar enkel voor onderzoek naar kansspelsverslaving en enkel via artikel 31m lid 5 Wok. Andere wijzen van gegevensverstrekking zouden een aanpassing van deze bepaling vereisen.

Borgen van medewerking

Uit het voorgaande blijkt dat het juridisch mogelijk is om rechtmatig gegevens op andere wijze en voor andere doeleinden te delen dan momenteel wordt gedaan, bijvoorbeeld centraal, koppelbaar en voor andere doeleinden dan kansspelsverslaving. Er kan echter niet zonder meer worden uitgegaan van vrijwillige medewerking van de gehele online kansspelsector aan dit soort nieuwe gegevensverstrekkingen. We

hebben drie routes geanalyseerd waarmee deze medewerking mogelijk kan worden gewaarborgd:

- **Route M1: zelfregulering.** Het is in theorie mogelijk om gebruik te maken van (hybride) zelfregulering om spelersgegevens beschikbaar te maken voor onderzoek. De sector zou zich vrijwillig kunnen committeren aan een te kiezen oplossingsrichting, aanvullend op de bestaande verplichting van art. 31m lid 5 Wok (pure zelfregulering), of zou dit kunnen doen bij wijze van invulling van die verplichting (hybride zelfregulering). Vanwege de conflicterende belangen is er bij beide vormen echter een risico dat de zelfregulering moeilijk en traag tot stand komt, zonder volledige dekking in de sector, en slecht wordt nageleefd. Zodra er geen volledige dekking is, bestaat het risico dat aanbieders die niet bij de regulering zijn aangesloten gaan adverteren met het feit dat zij geen spelersdata delen met onderzoekers. Dit zal waarschijnlijk ten koste gaan van de representativiteit van de gegevens die wél beschikbaar zijn voor onderzoek, en daarmee ook ten koste van de betrouwbaarheid van de onderzoeksresultaten. Met het oog op deze risico's lijkt zelfregulering hier niet het aangewezen instrument.
- **Route M2: aanpassing kansspelregelgeving.** Door de kansspelregelgeving te wijzigen kunnen aanbieders simpelweg worden verplicht om mee te werken aan de te kiezen oplossing. Een voor de hand liggende route is om de uitwerking van de huidige verplichting van artikel 31m lid 5 Wok in lagere regelgeving aan te passen, om zo beter tegemoet te komen aan de onderzoeksbehoefte.^a Zo kan medewerking worden gewaarborgd zonder dat een wetswijziging nodig is. De huidige verplichting strekt enkel tot het meewerken aan de verstrekking van gegevens ten behoeve van onderzoek naar *kansspelverslaving*. Onzes inziens staat de regelgeving er echter niet aan in de weg om gegevens die door aanbieders verplicht worden aangeleverd voor onderzoek naar kansspelverslaving, vervolgens ook te gebruiken voor onderzoek naar andere thema's (alleen in de restrictieve interpretatie zou dit niet kunnen). Daarnaast strekt de huidige verplichting enkel tot het verstrekken van *geanonimiseerde* gegevens, maar in de analyse van rechtmatigheid (§3.2) bleek dat het mogelijk is om met geanonimiseerde gegevens toch koppelbaarheid te realiseren. Wel zou het vereiste van anonimisering vermoedelijk ten koste gaan van het detailniveau van de gegevens: er zal randomisering en/of generalisering nodig zijn om te voorkomen dat individuen identificeerbaar zijn op basis van unieke combinaties in de data. Als een hoog detailniveau gewenst is, zou het daarom beter zijn de Wok zelf te

^a Voor sommige wijzigingen kan worden volstaan met aanpassing van de Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving. Dat aanbieders de gegevens verstrekken 'op verzoek van een onderzoeksinstituting' is echter bepaald in het Besluit wrvk, dus als een oplossing niet meer uit gaat van deze één-op-één-relatie dan zal het Besluit moeten worden aangepast.

wijzigen, zodat aanbieders verplicht kunnen worden om mee te werken aan de gewenste gegevensdeling, zelfs wanneer hiermee (gepseudonimiseerde) gegevens worden gedeeld die mogelijk kwalificeren als persoonsgegevens.

- **Route M3: verzameling door het CBS.** Een andere route is gebruik te maken van de bevoegdheden van het CBS. Bij AMvB kan het CBS de opdracht en bevoegdheid krijgen statistisch onderzoek naar kansspelgedrag uit te voeren en hiervoor gegevens op te vragen bij kansspel aanbidders. Aanbidders zijn dan verplicht deze gegevens te verstrekken. Het CBS kan de gegevens die het voor zijn eigen statistisch onderzoek heeft opgevraagd vervolgens ook beschikbaar maken voor onderzoekers via de CBS Microdata-omgeving.

Conclusie

Om onderzoek te faciliteren zal vermoedelijk gebruik moeten worden gemaakt van een wettelijke plicht. Hiervoor zien we drie opties:

1. **In de lagere regelgeving onder artikel 31m leden 5 en 6 van de Wok (Route R1 + M2).** Omdat lid 5 bepaalt dat de gegevens geanonimiseerd moeten zijn zal hier sprake zijn van een lager detailniveau van de gegevens (meer randomisering en generalisering), zodat onderzoekers de betrokkenen niet kunnen identificeren. Ook zou de afbakening tot onderzoek naar kansspelverslaving beperkend kunnen werken, afhankelijk van wetsinterpretatie en hoeveel ruimte de regering zich hier veroorlooft bij het uitoefenen van de gedelegeerde bevoegdheid.
2. **Het wijzigen van artikel 31m Wok** zodat volledig duidelijk is wat er wel en niet met de gegevens mag en moet gebeuren **(Route R1/R2 en M2)**. Dit kost echter aanzienlijk meer tijd dan het wijzigen van lagere regelgeving. Als de eis van anonimisering van gegevens wordt behouden is Route R1 van toepassing. De wetgever kan er echter ook voor kiezen om een eis als pseudonimisering of 'passende waarborgen' op te nemen in plaats van de eis van anonimisering. Dit biedt meer ruimte voor verschillende oplossingsrichtingen, omdat dan ook met persoonsgegevens kan worden gewerkt als dit noodzakelijk is (Route R2). Ook maakt dit een hoger detailniveau van de gegevens mogelijk.
3. **Het nemen van een AMvB op grond van de Wet CBS (Route R2 en M3)**, zodat het CBS de gegevens op kan vragen bij de aanbidders. Het CBS gebruikt de gegevens voor zijn eigen statistisch onderzoek en maakt ze daarnaast ook beschikbaar via de Microdata.

4 Oplossingsrichtingen

Om tot een logische set oplossingsrichtingen te komen en deze gestructureerd te kunnen vergelijken, is een analysekader nodig. Dit beschrijven we in §4.1. Op basis van het analysekader formuleren we vier oplossingsrichtingen, die we beschrijven in §4.2. Vervolgens vergelijken we de oplossingsrichtingen en beschouwen we de verschillen in §4.3.

4.1 Analysekader

Om tot oplossingsrichtingen te komen expliciteren we enerzijds de ontwerpdoelen (waar moet een oplossing voor worden geboden?) en anderzijds de randvoorwaarden (aan welke externe factoren moeten in alle gevallen worden voldaan?).

Ontwerpdoelen

In paragraaf 2.3 beschreven we, op basis van de verkenning van de onderzoeksbehoefte en huidige invulling daarvan, de volgende belangrijkste knelpunten bij het huidige proces. Deze kunnen worden vertaald naar ontwerpdoelen (ook wel *functionele eisen*). Deze ontwerpdoelen zijn de volgende (in willekeurige volgorde, en voorzien van letters zodat er verderop naar kan worden verwezen):

✓ A	Er kunnen achtergrondkenmerken worden gekoppeld aan de spelersgegevens;
✓ D	Er is toegang tot de data voor andere onderzoeks doelen dan preventie van kansspelverslaving;
✓ K	Spelersgegevens zijn koppelbaar op spelersniveau tussen aanbieders, waardoor een integraal beeld van spelersgedrag kan worden gevormd voor analyse;
✓ O	De Ksa kan zélf meer soorten onderzoek uitvoeren;
✓ P	Het proces voor datatoegang is beter ingericht voor zowel onderzoekers als aanbieders (eenvoudiger, minder tijdrovend/kostbaar en/of consistent).

De oplossingsrichtingen moeten erop zijn gericht om één of meerdere van deze knelpunten te adresseren. We spreken in dit kader bewust van oplossings**richtingen** en niet van *oplossingen*. Een oplossingsrichting zien we als een *geheel van technische, organisatorische en juridische deeloplossingen* dat (samen) tegemoet komt aan (een deel van) de genoemde ontwerpdoelen, binnen randvoorwaarden. In de praktijk bestaat er een sterke mate van afhankelijkheid tussen de onderliggend deeloplossingen: kies je voor een bepaalde technische invulling, dan is ook een bepaalde specifieke juridische

oplossing nodig, en soms moet eerst een juridische stap worden gezet voordat een volgende stap kan plaatsvinden. Daarnaast hoeven niet alle ontwerpdoelen op hetzelfde moment in tijd te worden gerealiseerd. Bij ontwikkeling van beleid kan ervoor worden gekozen om een oplossingsrichting ten dele, of gefaseerd in tijd te effectueren. Het is bijvoorbeeld denkbaar dat op korte termijn de deeloplossingen ter verbetering van het proces worden ingezet, terwijl de koppelbaarheid met achtergrondgegevens iets is wat pas op langere termijn wordt gerealiseerd. Het kan echter wel zo zijn dat de keuze voor een bepaalde oplossingsrichting bepalend is voor hoe dat in de toekomst dan kan gebeuren.

Randvoorwaarden

In de juridische analyse in hoofdstuk 3 is een aantal juridische knelpunten (en bijbehorende oplossingen) benoemd waaraan oplossingsrichtingen moeten voldoen:

- **Alle^a aanbieders werken mee aan het delen van data, en deze medewerking is juridisch voldoende^b geborgd.** In §3.2 hebben we de verschillende routes hiervoor uitgewerkt. Niet alle routes zijn van toepassing op alle oplossingsrichtingen. Van belang is dat ten minste één van de genoemde routes kan worden toegepast.
- **De datadeling is rechtmatig (te maken)^c.** In hoofdstuk 3.2 hebben we de verschillende routes hiervoor uitgewerkt. Niet alle routes zijn van toepassing op alle oplossingsrichtingen. Van belang is dat ten minste één van de genoemde routes kan worden toegepast.

Naast de bovengenoemde randvoorwaarden is het denkbaar dat er aanvullende criteria zijn die de ontwerpruimte afbakenen, bijvoorbeeld de omvang van het budget dat beschikbaar is voor het realiseren van datadeling. Dergelijke criteria zijn afhankelijk van de uiteindelijke politieke keuzes en dus nog niet (goed) vast te stellen in deze fase.

^a In theorie zou ook voor een oplossing kunnen worden gekozen waaraan niet *alle* aanbieders deelnemen; vanuit onderzoeksperspectief zou (afhankelijk van de voorziene analyse) wellicht te accepteren kunnen zijn dat 20% van de gegevens ontbreekt. We nemen echter aan dat het de bedoeling is dat *alle* aanbieders meewerken. Wanneer dit niet het geval zou zijn, is het aannemelijk dat er een selectie-effect plaatsvindt (de niet-aanwezige aanbieders vertegenwoordigen een specifieke groep spelers, en/of spelers gaan specifiek spelen bij aanbieders die geen data delen, die daar zelfs mee zouden kunnen adverteren). Vanuit perspectief van regelgeving en Ksa (als toezichthouder) is gelijke behandeling van alle aanbieders eveneens logisch.

^b De vraag wat 'voldoende' hier betekent, is aan de beleidsmaker. Als er voldoende vertrouwen is dat aanbieders bijvoorbeeld op vrijwillige basis meewerken, is geen of zeer lichte juridische borging nodig.

^c Zoals in hoofdstuk 3 ook al opgemerkt is, is er juridisch gezien geen 'absolute zekerheid' te geven over rechtmatigheid – uiteindelijk is dat aan de rechter. Preciezer is daarom de volgende formulering: de oplossingsrichting moet zodanig zijn ingericht dat het risico dat deze later niet rechtmatig blijkt, in de ogen van de beleidsmaker afdoende is gemitigeerd.

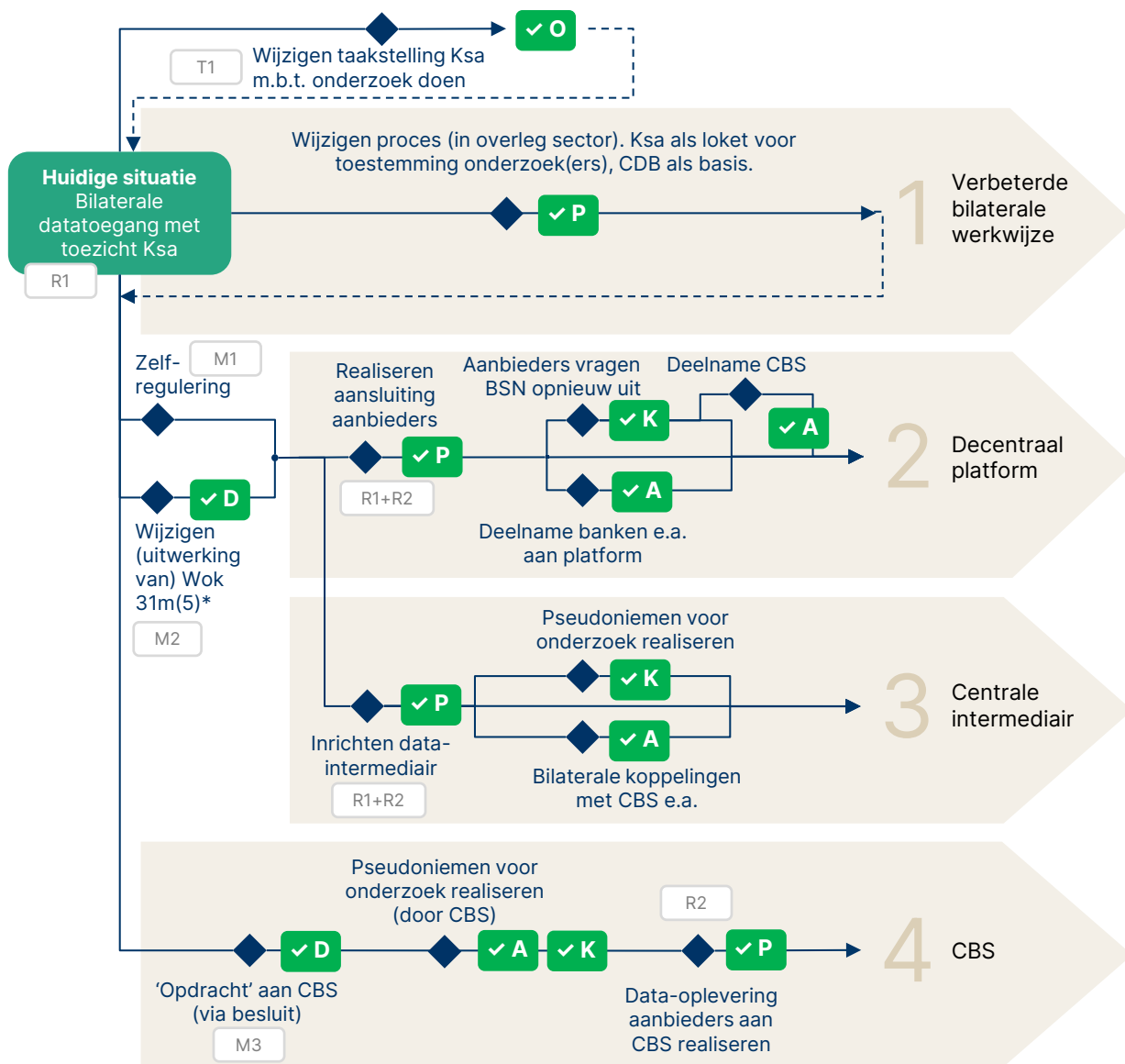
4.2 Oplossingsrichtingen

Onderstaande Figuur 1 toont schematisch de vier geformuleerde oplossingsrichtingen. Binnen deze oplossingsrichtingen zijn deeloplossingen aangegeven met een blauw ruitje. De groene rechthoeken geven aan wanneer een bepaald ontwerpdoel gerealiseerd is. De pijlen geven aan welke keuzes er gemaakt kunnen worden. Hieronder lichten we de vier oplossingsrichtingen op hoofdlijnen toe, waarna we ze één voor één in nader detail bespreken. De vier oplossingsrichtingen, zoals weergegeven in Figuur 1, zijn de volgende:

1. **Verbeterde bilaterale werkwijze.** In deze oplossingsrichting wordt het huidige proces, waarbij een onderzoek bilateraal gegevens opvraagt bij aanbieders, verbeterd. De verbetering bestaat uit het toepassen van twee deeloplossingen: (1) het aanpassen van de taakstelling van de Ksa, zodat zij zelf meer vormen van onderzoek kan uitvoeren, en/of (2) het vergroten van de rol van de Ksa in het proces, zodat het proces soepeler kan verlopen.
2. **Decentraal platform.** In deze oplossingsrichting wordt een nieuw platform ingericht waarop aanbieders data delen met onderzoekers. Dit gebeurt zoveel mogelijk decentraal. Met technische waarborgen wordt ervoor gezorgd dat partijen/personen die toegang hebben tot de data, en in principe beschikken over aanvullende informatie waarmee subjecten zouden kunnen worden geïdentificeerd, dit desondanks niet kunnen. Een voorbeeld is een (hypothetisch) platform waarop onderzoekers *queries* kunnen insturen, waarbij automatisch wordt vastgesteld of de uitkomsten voldoende geaggregeerd zijn. De onderzoeker 'ziet' in dit geval de data op individueel niveau niet, en alleen de geaggregeerde uitkomsten kunnen in handen komen van de organisaties/personen die over aanvullende informatie beschikken.

Binnen deze oplossingsrichting zijn verschillende technische invullingen denkbaar, die met name verschillen in de mate waarin er (alsnog) sprake is van een centrale entiteit die de benodigde onderliggende technische infrastructuur beheert. Met *secure multiparty computation* kan technisch worden gegarandeerd dat meerdere partijen die samen datadelen nooit ongeautoriseerd toegang kunnen krijgen tot gevoelige gegevens.

Bij het gebruik van een decentraal platform kan eventueel aanvullende data koppelbaar worden gemaakt, wanneer de bronhouder daarvan deze aanlevert op hetzelfde platform. Dat zou bijvoorbeeld het CBS kunnen zijn.



Legenda: gerealiseerde doelstellingen:

- ✓ K Spelers **koppelbaar** tussen aanbieders
- ✓ A **Achtergrondkenmerken** koppelbaar
- ✓ O Ksa kan zelf **onderzoek** doen
- ✓ D Datatoegang voor onderzoeken met **doel** breder dan kansspelverslaving
- ✓ P **Proces** voor datatoegang verbeterd voor onderzoekers en aanbieders

Gehanteerde juridische routes:

- | | | |
|---|------------------------------------|--|
| M1 Medewerking o.b.v. zelfregulering | R1 Verwerking o.b.v. anonimisering | T1 Onderzoek Ksa o.b.v. (gewijzigde) taakstelling en besluit Ksa |
| M2 Medewerking o.b.v. (ge-wijzigde) kansspelwetgeving | R2 Verwerking o.b.v. UAVG art. 24 | |
| M3 Medewerking o.b.v. wet CBS | | |

* Als geen persoonsgegevens hoeven te worden verwerkt kan worden volstaan met wijziging van de lagere regelgeving. Bij een restrictieve wetsinterpretatie is wijziging van 31m(5) zelf ook nodig t.b.v. onderzoek naar andere onderwerpen dan kansspelverslaving.

Figuur 1 Schematisch overzicht oplossingsrichtingen

3. **Centrale intermediair.** Een derde partij (een nieuwe of bestaande organisatie, zoals de Ksa) zou kunnen optreden als verzamel- en uitgiftepunt voor de gegevens. Dit is in de basis een sterk *organisatorisch* gerichte oplossing: de derde partij beoordeelt onderzoekers en onderzoeksverzoeken, vertaalt deze naar dataverzoeken aan de aanbieders, en zorgt ervoor dat de gegevens uiteindelijk worden gekoppeld en geanonimiseerd ter beschikking komen van de onderzoeker. Binnen deze oplossingsrichting worden naast organisatorische/juridische instrumenten voor het borgen van de privacy ook technische instrumenten ingezet (primair pseudonimisering).

Met sterke afspraken en pseudonimisering is het denkbaar dat onderzoekers zélf datasets ontvangen van de tussenpartij. Het is echter ook denkbaar dat de onderzoekers werken in een afgeschermd omgeving die de tussenpartij aanbiedt waarbij de tussenpartij een exportcontrole uitvoert (net zoals bij CBS Microdata en sommige implementaties van een decentraal platform).

4. **CBS.** In deze oplossingsrichting worden de data verzameld door het CBS, en beschikbaar gemaakt in de CBS Microdata-omgeving. Onderzoekers kunnen onder strenge voorwaarden op projectbasis toegang krijgen tot deze omgeving om onderzoek uit te voeren. Het is mogelijk deze richting zo in te richten dat de gegevens ook kunnen worden gekoppeld op persoonsniveau met andere datasets waarover het CBS beschikt, wat een belangrijk pluspunt kan zijn.

De vier oplossingsrichtingen kennen enige mate van overlap. Sommige deeloplossingen komen in meerdere oplossingsrichtingen voor. Daarnaast is er ook conceptueel enige overlap: zo is het CBS in oplossingsrichting 4 in feite ook een soort centrale intermediair. Bij de oplossingsrichting 'decentraal platform' zou de vraag kunnen worden gesteld in hoeverre de partij die het decentrale platform realiseert niet in feite óók een intermediair is. De reden om toch voor *deze* set oplossingsrichtingen te kiezen, is omdat zij fundamenteel verschillen in onder andere de wijze waarop aan de juridische randvoorwaarden wordt voldaan, of de keuze gevolgen heeft voor de mate en wijze waarop aan de ontwerpdoelen kan worden voldaan. Het CBS is in oplossingsrichting 4 weliswaar een soort intermediair, maar op basis van een geheel andere juridische basis dan de intermediair in oplossingsrichting 3. In oplossingsrichtingen 1 en 2 ligt het automatiseren van toegang tot de data primair bij de aanbieders, terwijl dat bij oplossingsrichtingen 3 en 4 nadrukkelijk bij een (ook inhoudelijk betrokken) centrale partij ligt.

De oplossingsrichtingen zijn wat ons betreft in principe wederzijds uitsluitend: het ligt niet voor de hand om één oplossingsrichting te kiezen en tegelijkertijd of later een andere (al is het niet onmogelijk). De eerste oplossingsrichting is daarop een uitzondering: het verbeteren van de huidige bilaterale werkwijze kan een tijdelijke, eerste stap zijn, waarna op de langere termijn een van de andere drie oplossingsrichtingen wordt ingeslagen. Daarnaast is het realiseren van verbreding van eigen onderzoek van

de Ksa (ontwerpdoel O) iets dat onafhankelijk van de overige richtingen kan worden opgelost (namelijk met het wijzigen van de taakstelling, zoals beschreven in §3.4).

4.2.1 Richting 1. Verbeteren bilaterale werkwijze

In deze oplossingsrichting wordt het huidige proces, waarbij een onderzoek bilateraal gegevens opvraagt bij aanbieders, verbeterd. Er zijn twee redenen om deze richting te verkennen. Allereerst beantwoordt het de vraag in hoeverre er met minimale aanpassingen en/of op korte termijn beter tegemoet kan worden gekomen aan de databehoefte. Anderzijds toont het uitwerken van deze richting de grenzen van de huidige opzet, waarmee een argument wordt gevormd voor het kiezen van een van de andere oplossingsrichtingen.

De oplossingsrichting bestaat uit het vergroten van de rol van de Ksa in het proces, zodat het proces soepeler kan verlopen. We zien hiervoor een aantal mogelijkheden:

- Een van de knelpunten die is gesignaleerd, is de complexe relatie tussen onderzoekers en aanbieders. De Ksa kan als partij *tussen* beiden gaan bemiddelen, waardoor de (gepercipieerde) afstand tussen onderzoekers en aanbieders wordt vergroot, maar er nog wel inhoudelijke informatie kan worden uitgewisseld.
- Een ander geconstateerd knelpunt betreft de consistentie van de beoordeling en behandeling van aanvragen door de verschillende aanbieders, en de mate waarin zij meewerken (ondanks de wettelijke verplichting). De Ksa zou kunnen bemiddelen door dataverzoeken van onderzoekers te *beoordelen*. Na goedkeuring wordt van de aanbieders verwacht dat zij meewerken. De Ksa zou verder kunnen bemiddelen bij vraagstukken over wat wel en niet gedeeld mag worden (zoals ten aanzien van de vereiste tot anonimisering).
- De aanbieders geven aan dat de variëteit aan verzoeken kan leiden tot veel werk, omdat iedere keer een nieuwe export moet worden gemaakt. Dit kan worden geadresseerd door het CDB-datamodel te gebruiken als 'standaard' voor alle verzoeken (onderzoekers kunnen dan wellicht nog variëren in de tijdsperiode die zij uitvragen, de selectie van aanbieders, et cetera) en daar alleen van af te wijken wanneer het echt nodig is. De Ksa is nu al verantwoordelijk voor het bepalen van het CDB-datamodel en zou hierbij dan ook de onderzoeksbehoefte kunnen betrekken.

Wanneer we kijken naar de andere ontwerpdoelen, dan stellen we vast dat deze niet (goed, binnen de randvoorwaarden) kunnen worden ingevuld op basis van alleen (een verbetering van) het huidige proces:

- **Koppelen van achtergrondkenmerken (A):** Dit zou vereisen dat aanbieders bilateraal aan onderzoekers datasets ter beschikking stellen, waarin de speler

(door een derde partij, zoals het CBS of een bank) geïdentificeerd kan worden. Dit zou betekenen dat de dataset per definitie niet meer geanonimiseerd is.^a

- **Koppelen van spelers tussen aanbieders (K).** De reden hiervoor is in beginsel dezelfde als hierboven: een dataset waarin een speler kan worden geïdentificeerd, is niet meer anoniem. Voor het koppelen tussen aanbieders zou echter ook kunnen worden gewerkt met een pseudoniem – zolang alle aanbieders dezelfde persoon dezelfde pseudoniem toekennen, zijn deze personen te koppelen tussen de datasets, maar niet identificeerbaar. In de praktijk is het echter zeer lastig om (zonder het inrichten van een specifieke oplossing hiervoor) geschikte pseudoniemen te hanteren. Aanbieders beschikken allereerst niet over een uniek kenmerk (BSN) om een persoon te identificeren (en op basis waarvan een goed pseudoniem kan worden berekend). Daarnaast is problematisch dat het pseudoniem dat vervolgens zou worden gebruikt, niet aanbiedersspecifiek en niet onderzoek-specifiek is. Dit zou betekenen dat een dataset voor onderzoek die in handen komt van een aanbieder, door deze aanbieder kan worden gekoppeld aan de eigen data, waarna deze spelers kan herkennen. Hetzelfde geldt voor onderzoekers die elkaars dataset inzien. Omdat bilateraal gegevens worden aangeleverd, is tot slot geen (goede) controle mogelijk om te voorkomen dat de *combinatie* van gegevens van verschillende aanbieders leidt tot een (individueel of groeps-)onthullingsrisico.
- **Toegang tot de data voor andere onderzoeksdoelen dan preventie van kansspelverslaving (D)** is niet te realiseren zonder een wijziging van artikel 31m van de Wok.

4.2.2 Richting 2. Decentraal platform

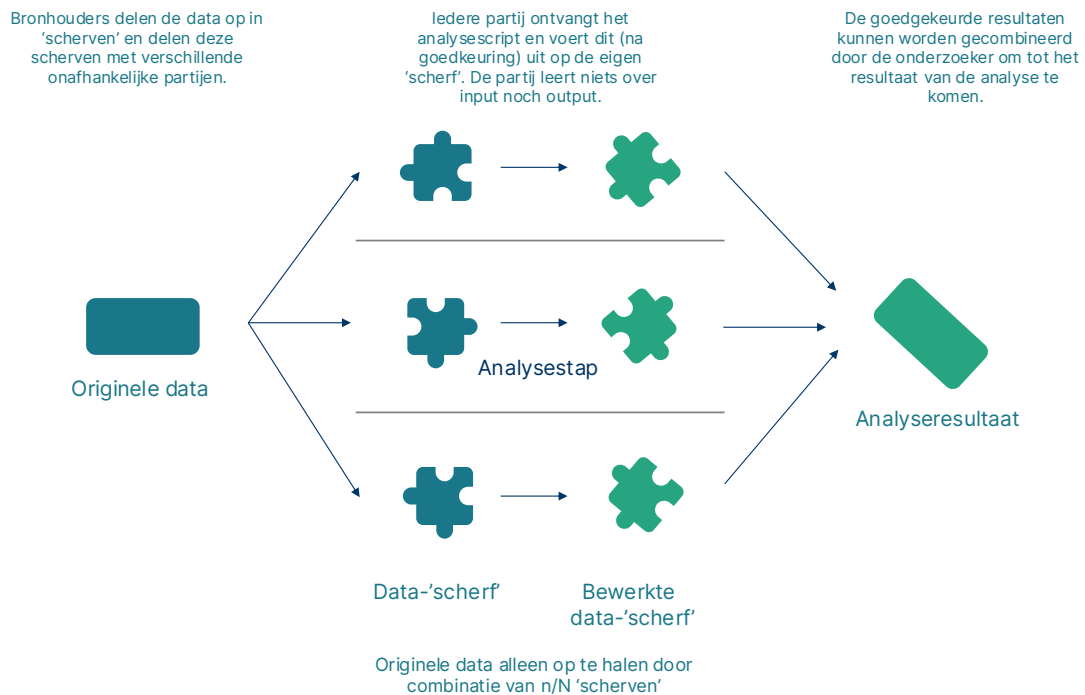
Het delen van data over speelgedrag (die koppelbaar is) op spelersniveau kan decentraal worden gerealiseerd met technische platforms. Een decentraal platform faciliteert daarbij bewerkingen op data zonder dat de onderzoeker de ruwe data kan inzien. Alleen het eindresultaat wordt vervolgens (wanneer de bewerking aan bepaalde eisen voldoet) vrijgegeven. Deze oplossingsrichting behelst het inrichten van een dergelijk platform en het daarop aansluiten van aanbieders, onderzoekers en eventueel derde partijen (als leveranciers van koppelbare data). Het op te richten platform staat los van de datadeling ten behoeve van toezicht (maar dezelfde datasets, en hetzelfde datamodel, zou uiteraard beschikbaar kunnen worden gemaakt via dit nieuwe platform, en het platform zou wellicht ook de datastroom voor toezicht kunnen faciliteren).

^a Daarnaast beschikken de aanbieders in deze opzet überhaupt niet over een kenmerk voor unieke identificatie van een persoon – het BSN wordt immers direct na registratie verwijderd. Omdat een aanbieder wel IBAN's opslaat, zou koppeling met bankgegevens in principe wel mogelijk moeten zijn.

Een 'eenvoudige' vorm van een dergelijke oplossing is er een waarbij onderzoekers een onderzoeksvraag (in de vorm van bijvoorbeeld een databasequery) zouden kunnen stellen aan een centrale intermediair of decentraal platform. Het platform voert de query uit wanneer de query en het resultaat aan bepaalde eisen voldoet (zoals bijvoorbeeld de door het CBS gehanteerde outputvereisten voor Microdata [CBS22]), en retourneert de resultaten aan de onderzoeker.

Hoewel dit een sterk technisch georiënteerde oplossing is, is er nog altijd sprake van een (niet-inhoudelijke) intermediair die het beschreven systeem en de bijbehorende dataverzameling en regels beheert. Door gebruik te maken van moderne cryptografische technieken kan een groter deel van deze aspecten worden geautomatiseerd. In de markt worden systemen aangeboden voor *secure multiparty computation* (SMPC). Een dergelijk systeem maakt gebruik van cryptografische principes om de ruwe data (die van meerdere partijen afkomstig kan zijn) op te splitsen in 'scherven'. Een data-scherf is op zichzelf onbruikbaar, omdat de originele ruwe data er niet van kan worden afgeleid. Een onderzoeker kan binnen de technische omgeving vervolgens bewerkingen uitvoeren op deze individuele scherven. Wanneer de onderzoeker resultaten wil exporteren, worden de (bewerkte) scherven gecombineerd en kan de uitkomst worden ingezien. Uiteraard worden de 'bewerkte scherven' alleen gedeeld wanneer de partij die de 'scherf' onder beheer heeft, toestemt met de analyse die de onderzoeker heeft uitgevoerd. Figuur 2 toont deze werkwijze schematisch. ^a

^a Een andere oplossing die in deze context werd genoemd tijdens de gevoerde gesprekken is *federated learning*. Hierbij kunnen modellen worden geschat op data die afkomstig is van verschillende bronhouders, zonder dat deze de ruwe data hoeven te delen. Deze techniek is specifiek relevant als wenselijk is om *machine learning* als analysemethode in te zetten (bijvoorbeeld voor het herkennen van onmatig speelgedrag).



Figuur 2 Werking van een decentraal platform voor secure multiparty computation op hoofdlijnen

In Nederland is TNO actief op het gebied van het ontwikkelen van deze en andere *privacy-enhancing technologies* (PET's) in haar PET-lab. [TNOPET] Het bedrijf Linksight is een spin-off van TNO dat op SMPC-gebaseerde oplossingen aanbiedt in onder andere de zorgsector. Linksight heeft daarnaast een pilot uitgevoerd waarbij ook het CBS aangesloten was als een van de datadelende partijen. [Lin25]

Een andere Nederlandse partij die een oplossing biedt op basis van SMPC is Roseman Labs. [Ros] De oplossing van Roseman Labs is geëvalueerd door de AIVD voor gebruik in overheidscontext (volgens het BSPA-kader) voor niet-staatsgeheime informatie. [AIVDRoseman] In de oplossing van Roseman Labs worden de 'datascherven' niet opgeslagen bij de verschillende bronhouders, maar verdeeld over servers die geheel gescheiden van elkaar worden beheerd (in principe door drie teams van Roseman Labs zelf). Op papier biedt dit een zeer hoge bescherming tegen datalekken, omdat een kwaadwillende alle drie deze omgevingen zou moeten binnendringen om tot de onderliggende data te kunnen komen.

SMPC-platforms zijn naast het faciliteren van structurele datadeling ook geschikt voor het ondersteunen van maatwerk-datadeling. Zo ondersteunt het platform van Roseman Labs het proces van data-aanvragen, waarbij aanbieders via een beveiligde webpagina gegevens kunnen importeren (die voorafgaand aan het uploaden direct worden versleuteld). Overigens is dit geen eigenschap die specifiek SMPC vereist – iets soortgelijks wordt door ZorgTTP aangeboden (zie richting 3). Het opleveren van maatwerkdata is handwerk voor de aanbieders en wellicht onwenselijk gezien de kosten die dit met zich meebrengt, en de beperkte meerwaarde (zie §2.1). Voor het structureel aanleveren van data volgens een vast datamodel moeten de aanbieders in ieder geval

een nieuw soort technische koppeling (ten opzichte van de bestaande koppeling voor aanlevering van CDB-data aan de Ksa) realiseren.

Bij toepassing van SMPC-gebaseerde oplossingen blijft het governancevraagstuk onverminderd bestaan. Afhankelijk van de wijze van inrichting zouden de data-aanbieders ofwel een inhoudelijke intermediair (zie richting 3) de 'toestemming' moeten geven voor het ontcijferen van resultaatdata. Daarbij kunnen kwalitatieve criteria spelen die niet automatisch kunnen worden bepaald. Bij SMPC speelt daarnaast dat de technische oplossing mogelijk beperkingen stelt aan het soort analyses dat kan worden uitgevoerd.

Een voordeel ten opzichte van richtingen 3 (intermediair) en 4 (CBS) is dat er in mindere mate of geen sprake is van centralisatie van gegevens. In het huidige klimaat waarin in toenemende mate gegevens over individuen worden verzameld, geanalyseerd en ook lekken (met alle gevolgen van dien), kan dat een relevante overweging zijn, zeker gelet op het feit dat het hier gaat om gegevens in relatie tot (mogelijke) verslaving.

4.2.3 Richting 3. Centrale intermediair

In deze oplossingsrichting faciliteert een centrale intermediaire partij de datadeling: de aanbieders leveren aan de intermediair, en de intermediair levert aan onderzoekers.

De eerste helft van deze opzet bestaat op dit moment in feite al bij de Ksa: alle aanbieders leveren op dit moment al CDB-gegevens aan de Ksa. De Ksa kan deze gegevens echter niet zomaar doorleveren aan onderzoekers (laat staan zelf voor alle soorten onderzoek gebruiken) – hiervoor zou een wijziging in de taakstelling nodig zijn. Mogelijk is ook een wetswijziging nodig om medewerking te borgen (ook als de Ksa niet de intermediair is).

Daarnaast zijn de spelers in de CDB gepseudonimiseerd, met door de individuele aanbieders gekozen pseudoniemen. De Ksa kan dus geen koppeling realiseren op spelerniveau tussen de datasets van de aanbieders. Om deze reden moeten aanbieders een ander pseudoniem gaan gebruiken in de data. Hieronder gaan we in op mogelijkheden voor het genereren van dergelijke pseudoniemen, waarbij mogelijk kan worden aangesloten bij de bestaande werkwijze van het CRUKS-register.

Om de doelstellingen van het mogelijke onderzoek te verbreden (ontwerpvereiste D) moet, net als in oplossingsrichting 2, artikel 31m van de Wok worden aangepast. Dit is geen vereiste voor het voortzetten van de oplossingsrichting (de intermediair kan

uiteraard alleen delen ten behoeve van onderzoek naar preventie van kansspelverslaving).^a

Om de Ksa zélf de mogelijkheid te geven breder onderzoek te doen, moet de taakstelling van de Ksa worden gewijzigd. Dit is ook aan de orde wanneer de Ksa de rol van intermediair toebedeeld krijgt, en wanneer zij een rol krijgt bij de toewijzing van pseudoniemen. Hieronder gaan we in op de vraag wie de intermediair kan worden, en hoe pseudoniemen voor spelers kunnen worden gegenereerd.

Voor het koppelen van achtergrondgegevens zijn bilaterale afspraken nodig tussen de intermediair en de bronhouder van de te koppelen gegevens. Het is in deze opzet iets lastiger om gegevens van het CBS te kunnen koppelen dan in oplossingsrichting 4: de gegevens van het CBS mogen de CBS-omgeving niet verlaten, dus de gegevens over spelersgedrag zouden moeten worden 'geïmporteerd' in de CBS-omgeving. Daarbij moet dan op basis van pseudoniemen van de intermediair kunnen worden gekoppeld met de CBS-pseudoniem (RINpersoon).

Wie wordt de intermediair?

Een relevante vraag in deze oplossingsrichting is uiteraard welke organisatie de rol van intermediair gaat vervullen. Er is een aantal (in onze ogen, op basis van de gevoerde gesprekken) logische keuzes. We merken hierbij op dat de genoemde organisaties zich op dit moment niet *formeel* committeren aan het kunnen of willen vervullen van de rol van intermediair.

- **De Ksa.** De aanbieders hebben uiteraard al nauw contact met de Ksa, en delen uiteraard al data ten behoeve van de CDB's. Een knelpunt bij de keuze voor de Ksa is dat de Ksa niet tot taak heeft om onderzoeksdata te verzamelen en te delen (de taken zijn gedefinieerd in art. 33b Wok). De Ksa is op dit moment betrokken bij aanvragen voor onderzoeksdata bij individuele aanbieders (via artikel 31m Wok). Formeel gezien vraagt een onderzoeker de data rechtstreeks op bij de aanbieder; de Ksa is uitsluitend betrokken om te monitoren of dit proces goed verloopt, en om eventueel te bemiddelen bij verzoeken.
- **Een bestaande intermediair.** Gedacht kan worden aan Stichting Informatievoorziening Zorg (IVZ), die de uitwisseling van data over verslaving via het LADIS uitvoert. Deze taak van het IVZ is wettelijk vastgelegd. Het IVZ besteedt het uitwisselen van de data voor LADIS (en in het bijzonder ook het uitgeven van de pseudoniemen aan dataleveranciers enerzijds en datagebruikers anderzijds) uit aan stichting ZorgTTP. De rol van het IVZ bij LADIS is inhoudelijk; zo worden door IVZ ook bepaalde analyses en duidingen uitgevoerd op de data.

^a Wanneer de 'strengere' interpretatie van Wok 31m ten aanzien van anonimisering geldt (als breed geldende vereiste en niet als smalle verplichting) dan is wijziging van 31m ook nodig.

Verzoeken voor (maatwerk)data worden in principe ook door het LADIS behandeld, maar zijn op dit moment niet aan de orde vanwege onzekerheid over de wettelijke basis hiervoor.^a

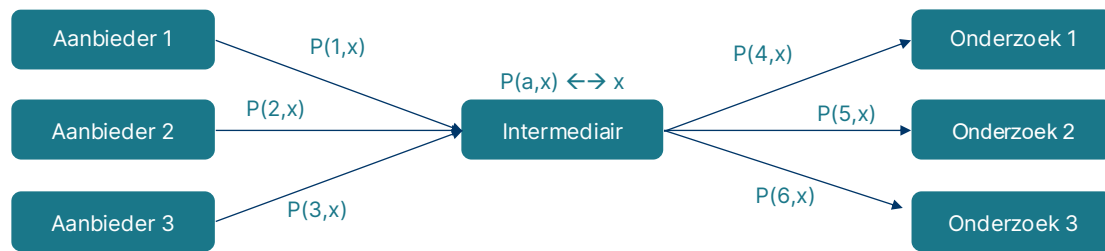
- **Een nieuw op te richten intermediair.** Deze intermediair zou (net als het IVZ) de uitwisseling zelf kunnen overlaten aan een partij als ZorgTTP. De rol van de intermediair is om te beslissen over wijzigingen aan het datamodel en om te bemiddelen tussen onderzoekers en aanbieders bij vragen over de data. De rol van nieuwe intermediair kan uiteraard worden vervuld door een bestaande organisatie, zoals een bestaande branchevereniging.
- **Het CBS.** Wanneer het CBS de intermediair wordt, is er sprake van oplossingsrichting 4 (zie verderop). Dit is een separate oplossingsrichting, omdat hiervoor een andere juridische route moet worden bewandeld.

Voor het realiseren van de benodigde technische infrastructuur kan gebruik worden gemaakt van de diensten van onder andere ZorgTTP (pseudonimisering en uitwisseling). Naast het LADIS faciliteert ZorgTTP ook data-uitwisselingen voor onder andere Nivel (zorgregistraties eerstelijns) en de NZa (DIS-registratie). Een andere relevante leverancier is SURF; deze biedt met SANE (*Secure ANalysis Environment*) een technische omgeving waarin analyse kan worden uitgevoerd door onderzoekers, maar de bronhouder volledig in controle blijft over de data. [SUR] De technische platforms uit oplossingsrichting 2 kunnen ook worden gebruikt, waarbij het essentiële verschil met die oplossingsrichting is dat een intermediair (en niet de individuele dataleveranciers) beslissen over onderzoeksaanvragen en data-exports.

Uitgifte van pseudoniemen

In een oplossing waarin een intermediair gegevens verzamelt van de verschillende aanbieders, koppelt de intermediair de gegevens op spelerniveau tussen de aanbieders. Het is daarbij wenselijk om zowel aan de aanbiederskant als aan de onderzoekerskant te werken met gebruikersspecifieke (en aan de onderzoekerskant mogelijk zelfs onderzoeksspecifieke) pseudoniemen. Figuur 3 toont deze werkwijze schematisch.

^a Als gevolg van de invoering van de AVG is de afgelopen jaren gewerkt aan de wettelijke basis van de gegevensverwerking voor het LADIS, maar niet voor extern gebruik van deze data



Figuur 3 Dubbelblinde en aanbieder/onderzoek-specifieke pseudoniemen

In Figuur 3 stelt $P(a, x)$ een pseudoniem voor dat een speler binnen dataset a uniek identificeert op basis van onderliggend persoonsgebonden (uniek) gegeven x (de parameter a zou een ‘toepassingsdomein’ kunnen worden genoemd van het pseudoniem). Kennis van $P(a, x)$ geeft geen enkele kennis over x en geen enkele kennis over $P(b, x)$. Een aanbieder weet dus niet welk pseudoniem een andere aanbieder hanteert voor een bepaalde speler. Een onderzoeker kan eveneens geen pseudoniemen koppelen tussen verschillende onderzoeksdatasets. De intermediair moet uiteraard in staat zijn om $P(a, x)$ te genereren gegeven een bepaalde x . Daarnaast is het nodig om de $P(a, x)$ te kunnen ontcijferen naar de onderliggende x , zodat de intermediair de data van verschillende aanbieders kan koppelen op x .

De x in Figuur 3 kan een gegeven als het BSN zijn. Het heeft echter sterk de voorkeur om het BSN niet te verwerken, vanwege het feit dat dit een bijzonder persoonsgegeven betreft. Het is technisch gezien ook niet *nodig* om het BSN te verwerken om tot koppelbaarheid op persoonsniveau van de gedragsdata te komen (en omwille van minimalisatie van verwerking van persoonsgegevens dus ook niet de bedoeling). Wanneer voor x een gegeven wordt gebruikt dat niet (direct) tot een persoon herleidbaar is (zoals bijvoorbeeld een *salted hash*^a van een gegeven als het BSN), dan heeft ook de intermediair geen mogelijkheid om gegevens tot een persoon te herleiden. In het CRUKS wordt deze werkwijze toegepast (zie hieronder).

Polymorfe pseudoniemen en het BSNk PP

Een geavanceerde(re) implementatie van bovenstaande zijn zogenoemde *polymorfe pseudoniemen*. Een polymorf pseudoniem is ontvangerspecifiek, en wordt gegenereerd

^a Een *hash* is een wiskundige ‘vingerafdruk’ die kan worden berekend over een gegeven (zoals een BSN-nummer). Voor ieder gegeven is er precies één hashwaarde (en de kans dat twee gegevens dezelfde hashwaarde hebben is astronomisch klein). Het is wiskundig nagenoeg onmogelijk om op basis van een hashwaarde de originele gegevens te herleiden. Hashes zijn vanwege deze eigenschap goed bruikbaar voor pseudonimisering, maar geen panacee: iemand die beschikt over een BSN kan eenvoudig zelf de bijbehorende hash uitrekenen en controleren of die persoon in de dataset voorkomt. Erger nog: iemand die beschikt over *alle mogelijke* invoergegevens (in het voorbeeld: alle mogelijke BSN’s) kan ook *alle bijbehorende* hashes genereren (een zogenoemde ‘rainbow table’ maken) en zo alsnog een hash herleiden tot een speler. Door aanvullende informatie toe te voegen aan de gehashte gegevens kan dit worden gemitigeerd. Zolang die aanvullende informatie geheim wordt gehouden, kan een aanvalleur noch individuele hashes voor een specifieke BSN, noch een ‘rainbow table’ genereren.

op basis van een *polymorfe identiteit* die door een centrale tussenpartij wordt beheerd. [Log24 p. 48, 6.2.2] Het verschil met de hierboven beschreven werkwijze is dat specifiek voor dit doel ontwikkelde cryptografische technieken worden gebruikt. [Log24]

Voor (overheids)organisaties die een betrouwbare digitale identiteit van een burger moeten verwerken, maar niet het BSN mogen verwerken, is binnen de Nederlandse Overheid Referentie Architectuur (NORA) het BSN-koppelregister Polymorfe Pseudonimisering (BSNk PP) beschikbaar. [ICT22] Dit koppelregister geeft op basis van polymorfe identiteiten polymorfe pseudoniemen uit, en is beschikbaar voor organisaties die (zoals in de voorliggende casus) optreden als makelaar ten behoeve van een publieke taak. Het BSNk PP vereist echter dat de betreffende persoon zich aanmeldt via een eIDAS-authenticatiedienst [Log24 p. 25, 5.22], iets wat in het voorliggende vraagstuk op dit moment niet aan de orde is. Om het BSNk PP hier te kunnen gebruiken, zou een speler dus (bij registratie of in ieder geval eenmalig per aanbieder) via eIDAS moeten authenticeren.

Bestaande pseudoniemen: CRUKS en overkoepelende speellimieten (SpeeLR)

CRUKS en het systeem voor overkoepelende speellimieten zijn beide systemen (bestaand en in ontwikkeling) die aanbiederoverstijgend spelers koppelen, op basis van pseudonimisering. De voor het CRUKS gehanteerde pseudoniemen werken volgens de hierboven en in Figuur 3 toegelichte methode en zijn dus relevant om nader te bekijken.

CRUKS

CRUKS is een afkorting voor *Centraal Register Uitsluiting Kansspelen*. Een persoon kan in het CRUKS worden geregistreerd of zichzelf registreren voor een bepaalde tijd. Wanneer deze persoon binnen deze termijn probeert in te loggen op een online kansspelplatform of een fysieke goklocatie poogt te bezoeken, wordt toegang door de kansspelaanbieder geweigerd wanneer de persoon voorkomt in het CRUKS. Bij de eerste controle stuurt de (online) aanbieder persoonsgegevens (zoals BSN, naam en geboortedatum) naar het CRUKS. CRUKS verifieert deze gegevens via de Gemeentelijke Basisadministratie Persoonsgegevens (GBA) en genereert vervolgens een unieke interne CRUKS-code (concreet een 'salted hash'^a van een aantal gegevens uit het GBA). Aan zijde van CRUKS worden vervolgens de oorspronkelijke gegevens verwijderd. De interne code wordt vervolgens versleuteld tot een externe aanbiederspecifieke CRUKS-code, die teruggaat naar de aanbieder. Wanneer dezelfde speler zich registreert bij een andere aanbieder is de interne CRUKS-code hetzelfde, maar krijgt de andere aanbieder een eigen aanbiederspecifieke externe CRUKS-code voor deze speler.

Bij latere inlogpogingen van de speler stuurt de aanbieder de opgeslagen externe CRUKS-code naar CRUKS. CRUKS kan de aanbiederspecifieke CRUKS-code

ontcijferen tot de interne CRUKS-code die in het register staat. Als dat zo is geeft CRUKS dit door aan de aanbieder en wordt de speler geweigerd.

Herbruikbaarheid van de CRUKS-pseudoniemen

De aanbiederspecifieke CRUKS-codes zouden door de aanbieders kunnen worden opgenomen in de op te leveren data. Het CRUKS beschikt over een sleutel om deze aanbiederspecifieke CRUKS-codes te ontcijferen tot een interne CRUKS-code, die aanbiederoverstijgend een speler identificeert. Deze CRUKS-code zou vervolgens kunnen worden gebruikt voor de koppeling. Daarna kan de speler met een nieuwe code (bijvoorbeeld een hash van de interne CRUKS-code of een heel nieuwe code) worden geïdentificeerd in de uiteindelijke dataset. Door de gehanteerde werkwijze wordt in geen van deze stappen het BSN of een ander direct herleidbaar pseudoniem verwerkt, noch door de aanbieders, noch de centrale partij.

De CRUKS-code is echter nooit bedoeld geweest voor pseudonimisering van spelers in onderzoeksdata. Het feit dat de code in de data die de aanbieders delen mogelijk ook iets zegt over registratie in het CRUKS is niet nodig en maakt het pseudoniem nodeloos gevoelig.

Herbruikbaarheid van de *werkwijze* voor CRUKS-pseudoniemen

De oplossing hiervoor zou zijn om, volgens dezelfde *methode* als CRUKS (en wellicht zelfs *bij* CRUKS) een tweede CRUKS-code ('Onderzoekscore') op te vragen om als pseudoniem te gebruiken. Het voordeel hiervan is dat bestaande systemen kunnen worden hergebruikt, en er niets nieuws hoeft te worden ontwikkeld. In plaats van één CRUKS-code vragen de aanbieders op soortgelijke wijze *nóg* een code op (een 'Onderzoekscore').

De CRUKS-code wordt gegenereerd bij de eerste registratie van een speler op een online platform, op basis van persoonsgegevens die alleen op *dát* moment beschikbaar zijn en daarna worden vernietigd door de aanbieder. Het is daarom niet mogelijk voor een aanbieder om een 'tweede CRUKS-code' te genereren voor bestaande spelers. Oplossingen hiervoor zijn:

1. De aanbieders de spelers laten vragen het BSN nog eenmaal in te vullen, zodat de aanbieder een tweede CRUKS-code kan genereren.
2. Eenmalig de 'tweede CRUKS-codes' te laten genereren en opleveren aan de aanbieders (het CRUKS kan dit doen, omdat het over de sleutels beschikt om de CRUKS-codes te ontcijferen. Het vereist dat daarna gegenereerde 'tweede CRUKS-codes' op basis van dezelfde methode en sleutel worden gegenereerd)
3. De bestaande CRUKS-code gebruiken als pseudoniem, en voor nieuwe spelers een nieuwe Onderzoekscore genereren en gebruiken. Het hergebruiken van de CRUKS-code in onderzoeksdata lijkt echter ongewenst.
4. Alleen gegevens van nieuw geregistreerde spelers in de onderzoeksdataset opnemen.

Overkoepelende speellimieten (SpeeLR)

Parallel aan dit onderzoek vindt een verkenning plaats naar de ontwikkeling van een systeem voor overkoepelende speellimieten. Zo'n systeem moet voorkomen dat spelers bij het bereiken van hun limiet bij de ene aanbieder uitwijken naar een andere aanbieder waar zij hun limiet nog niet bereikt hebben. Overkoepelende speellimieten kunnen daarmee een belangrijke rol spelen in het beperken van onmatig speelgedrag, vroegsignalering en tijdige interventie.^a

Het principe van overkoepelende speellimieten verschilt van dat van CRUKS. Waar CRUKS een lijst bijhoudt van uitgesloten spelers, registreert het systeem voor overkoepelende speellimieten hoeveel een speler mag inzetten. Fundamenteel gezien zou een aanpak zoals bij CRUKS voor wat betreft de pseudoniemen ook hier kunnen werken: in plaats van het registreren of een bepaalde CRUKS-code wél of níet geregistreerd staat, moet SpeeLR voor iedere 'SpeeLR-code' een getal registreren.

Voor het realiseren van overkoepelende speellimieten is een wijziging van de AMvB en waarschijnlijk de wet nodig.^b De exacte technische en organisatorische invulling van het systeem wordt nader onderzocht en is onder andere afhankelijk van de uiteindelijke beleidskeuzes die worden gemaakt. Het ligt echter ons inziens voor de hand om een eventueel vraagstuk over pseudoniemen voor onderzoek te laten 'meelopen' met het vraagstuk voor de pseudoniemen voor SpeeLR vanwege de grote overlap.

Wanneer (zoals hierboven voorgesteld) aan spelers opnieuw het BSN moet worden gevraagd om een pseudoniem te kunnen genereren voor SpeeLR, dan zou ook direct van de gelegenheid gebruik kunnen worden gemaakt om een onderzoekspseudoniem te genereren. Spelers hoeven dan niet twee keer te worden gevraagd om een BSN, de aanbieder hoeft het BSN niet twee keer te verwerken, en krijgt waarschijnlijk minder vragen over de BSN-uitvraag. Het zou ook voor de hand liggen om hiervoor aan te sluiten bij BSNk PP.

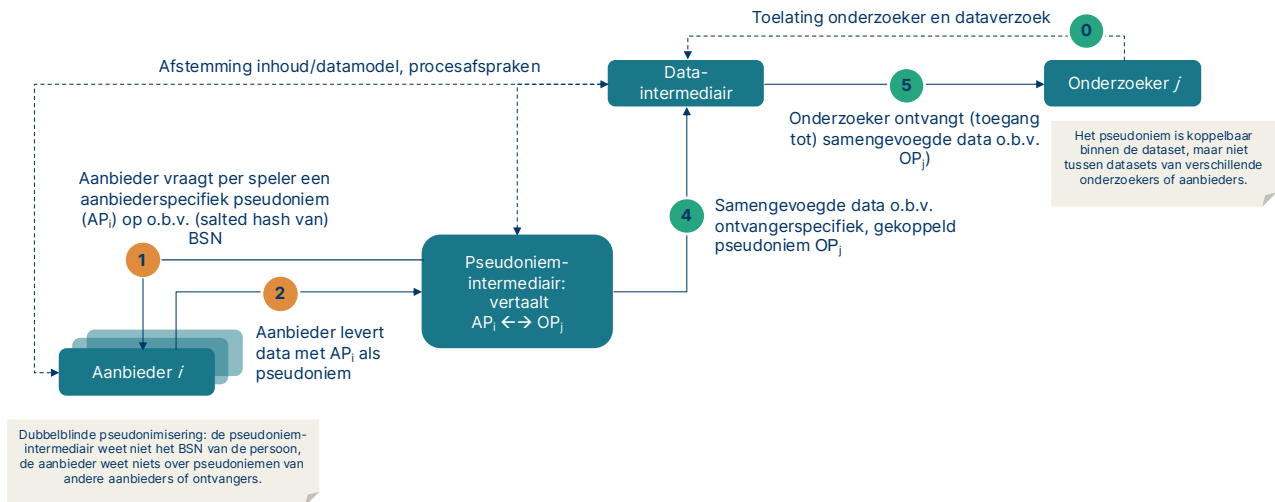
Een enkele of gesplitste intermediair

Wanneer wel met een **x** wordt gewerkt die herleidbaar is tot een persoon, kan de rol van de intermediair worden gesplitst: één rol betreft het uitgeven van pseudoniemen en koppelen van de data, de andere het delen van de data en het inhoudelijke aspect (afstemmen over het datamodel, et cetera). Een dergelijke splitsing kan ook om andere redenen zinvol zijn. Zo zou de Ksa de rol van 'deler' op zich kunnen nemen (zoals IVZ dit doet voor LADIS), en de pseudonimisering en koppeling kunnen uitbesteden aan een

^a Eerste Kamer, vergaderjaar 2023-2024, 33996, Y

^b Tweede Kamer, vergaderjaar 2023-2024, 24557, nr. 220

partij als ZorgTTP (ibid.). Figuur 4 toont de opzet met een 'gesplitste' intermediair schematisch.



Figuur 4 Mogelijke organisatorische opzet met een gesplitste intermediair (een 'koppelaar' en een 'deler').

4.2.4 Richting 4. CBS Microdata

In deze paragraaf bespreken we een aantal aspecten van de oplossingsrichting CBS Microdata die specifiek onderscheidend zijn ten opzichte van de andere oplossingsrichtingen.

Toegangsbeperkingen

Om toegang te krijgen tot de CBS Microdata-omgeving en om hierin een onderzoek te kunnen uitvoeren, moeten zowel de uitvoerende organisatie, de onderzoeker als het onderzoek aan specifieke eisen voldoen. Hieronder lichten we deze eisen toe en wat dit betekent voor de inclusiviteit van toegang tot de data.

Vereisten aan de organisatie

Toegang tot CBS Microdata is voorbehouden aan (1) Nederlandse universiteiten, (2) bij wet ingestelde organisaties of instellingen voor wetenschappelijk onderzoek, (3) bureaus voor statistiek binnen de EU, en (4) onderzoeksafdelingen van ministeries en andere diensten, organisaties en instellingen. In deze laatste categorie kunnen ook niet-publieke organisaties vallen, als deze voldoen aan de volgende voorwaarden: (1) de hoofdactiviteit is het verrichten van statistisch of wetenschappelijk onderzoek, of er is een 'statistische enclave'^a ingericht die deze activiteiten uitvoert, (2) de resultaten

^a Een statistische enclave is volgens het CBS een 'aparte afdeling die volledig gericht is op het verrichten van onderzoek en die onder eigen leiding staat, onafhankelijk van de overige activiteiten van uw instelling'. [26]

worden openbaar gemaakt, (3) de organisatie heeft rechtspersoonlijkheid^a en is gevestigd binnen de EER. [CBSRA].

Vereisten aan de onderzoekers

Alle individuele onderzoekers die (voor hun instelling) toegang krijgen tot CBS Microdata, moeten vooraf een examen afleggen, en worden periodiek gecontroleerd op hun kennis van de regels voor het werken met Microdata. Daarnaast moeten onderzoekers (vanaf eigen apparatuur op afstand) werken in de omgeving van CBS Microdata, waarbij diverse beveiligingsmaatregelen in acht moeten worden genomen. Het is niet te verwachten dat deze eisen noemenswaardige knelpunten opleveren voor de doelgroep.

Vereisten aan het onderzoek

Ieder onderzoek dat binnen CBS Microdata wordt uitgevoerd, wordt vooraf getoetst door een onderzoekscommissie. De onderzoeker (in dienst van een gemachtigde organisatie) dient hiervoor een projectaanvraag in bij het CBS. Het onderzoek moet voldoen aan de volgende eisen: [CBSRA]

1. Het onderzoeksdoel moet vooraf helder zijn, in verband met de doelbinding die de AVG vereist. Exploratief onderzoek is evenwel een toegestaan doel.
2. Het *doel* van het onderzoek moet statistisch van aard zijn. Het mag dus niet voor administratieve, gerechtelijke of fiscale doeleinden worden gebruikt. Het onderzoek mag niet tot doel hebben om op niveau van individuele personen of organisaties te controleren.
3. De resultaten van het onderzoek moeten openbaar worden gemaakt, en vooraf moet worden bepaald hoe dit gebeurt.
4. Het moet aannemelijk zijn dat de onderzoeksvraag kan worden beantwoord met CBS Microdata. Als de vraag ook buiten de Microdataomgeving kan worden beantwoord, dan moet dit dus buiten de omgeving gebeuren. Daarnaast moet vooraf worden vastgesteld dat de benodigde data (en specifieke variabelen) bestaan in de datacatalogus van het CBS.
5. Als aanvullende datasets worden ingebracht ("geïmporteerd") in de Microdataomgeving moeten deze rechtmatig zijn verkregen, en moet de onderzoeksinstelling deze data ook rechtmatig mógen importeren en gebruiken voor het onderzoek.
6. De inhoud en de methode van het onderzoek blijven de verantwoordelijkheid van de onderzoeker – met andere woorden, de naam van het CBS wordt niet aan het onderzoek verbonden.

^a Dit betekent voor in Nederland gevestigde organisaties dat dit een B.V., N.V., stichting, vereniging, coöperatie, onderlinge waarborgmaatschappij of kerkgenootschap moet zijn. Nederlandse eenmanszaken, V.O.F.'s, maatschappen, commanditaire vennootschappen en rederijen zijn uitgesloten. [26]

Voordat (resultaat)gegevens uit de Microdataomgeving mogen worden geëxporteerd, voert het CBS een controle uit van de te exporteren data. Het CBS controleert hierbij primair op het onthullingsrisico – kunnen gegevens worden herleid tot individuele personen of organisaties? Bij de controle wordt gekeken naar de aangeboden dataset in isolatie (en dus niet naar onthullingsrisico's door combinatie van separate exports). Naast de te exporteren data wordt ook de gehanteerde 'syntax' (achterliggende scripts) in overweging genomen. De vereisten aan te exporteren data zijn door het CBS uitvoerig beschreven. [CBS22] De regels beschrijven onder andere het aantal vrijheidsgraden in modellen en een minimale celvulling. Wellicht triviaal maar belangrijk is dat ook rechtstreekse export van microdata wordt uitgesloten.

Conclusie

Bovenstaande betekent dat CBS Microdata toegankelijk is voor wetenschappelijk onderzoek (mits uitgevoerd door een universiteit) en maatschappelijk onderzoek (uitgevoerd door maatschappelijke organisaties, mits zij daarvoor een enclave inrichten en over een rechtspersoonlijkheid beschikken). Over het algemeen zullen de eisen die op onderzoeksniveau worden gesteld geen noemenswaardige beperking vormen voor deze groep. Een groot deel van de organisaties is zelfs mogelijk al gemachtigd voor toegang tot CBS Microdata.

Voor commerciële onderzoeksbureaus is toegang eveneens mogelijk, voor zover de organisatie de juiste rechtsvorm heeft en daadwerkelijk een onderzoeksbureau is. Voor deze groep vormen de eisen aan het onderzoek waarschijnlijk alleen een mogelijk knelpunt wanneer het onderzoek niet in opdracht van overheden of maatschappelijke instituten wordt uitgevoerd.

Voor *aanbieders* is toegang in principe ook mogelijk, mits zij een enclave voor onderzoek inrichten (rechtspersoonlijkheid mag worden aangenomen) en binnen de EER gevestigd zijn. Het is echter aannemelijk dat de aanbieders voor dergelijk onderzoek een doel zullen hebben dat niet past binnen de vereisten van het CBS. Ook het vereiste om onderzoeksresultaten te publiceren kan voor de aanbieders een beperking vormen. Het is uiteraard denkbaar dat aanbieders wetenschappelijk/maatschappelijk onderzoek (laten) uitvoeren dat wél binnen de doelvereisten van het CBS past. Niet uit te sluiten is dat dergelijk onderzoek ook een secundair (commercieel) doel dient.

Relevant is tot slot dat (zo is althans de ervaring van de auteurs van dit onderzoek) de werkwijze van CBS in hoge mate *maatwerk* is. De beoordeling van onderzoeken en toelating is mensenwerk en biedt ruimte voor afstemming en eventueel het aanpassen van onderzoeksvoorstellen.

Verzamelen van de gegevens

Het CBS verzamelt de gegevens rechtstreeks bij de aanbieders. In §3.3 is beschreven dat het CBS gebruik zou kunnen maken van de bij wet bepaalde bevoegdheid om gegevens rechtstreeks te verzamelen bij aanbieders. Uiteraard zou ook een andere

(wettelijke) verplichting kunnen worden geformuleerd die de aanbieders verplicht de data aan het CBS op te leveren. In deze situatie is eenzelfde oplossing nodig voor wat betreft de pseudoniemen, als de aanbieder het BSN niet kan, wil of mag opslaan.

Koppelbaarheid met achtergrondkenmerken

Binnen de CBS Microdataomgeving kunnen datasets van het CBS op persoons- of organisatieniveau worden gekoppeld. Het CBS gebruikt hiervoor een pseudoniem ("RINpersoon") die is afgeleid van het BSN. Wanneer de gegevens over speelgedrag op spelerniveau zijn voorzien van deze RINpersoon-identificatie kunnen deze dus worden gekoppeld met een grote verscheidenheid aan achtergrondkenmerken, variërend van gegevens over werk en inkomen tot gezondheid. [CBSMicrodata]

Omdat alleen het CBS beschikt over de 'sleutel' om het BSN te vertalen naar RINpersoon, kan de spelersdata alleen worden voorzien van RINpersoon als het CBS de data over speelgedrag kan koppelen op BSN. Wanneer wordt gewerkt op basis van het CDB-datamodel, zou dit vereisen dat de (aanbiedersspecifieke) pseudoniemen die door de aanbieders worden gehanteerd in de aangeleverde data zijn afgeleid van het BSN, en (alleen) een tussenpartij (vergelijkbaar met CRUKS) deze kan ontcijferen tot het BSN. Deze partij zal dan met het CBS een manier moeten afstemmen om tot koppeling met het BSN te kunnen komen. Wanneer gebruik wordt gemaakt van de CBS-bevoegdheid voor het rechtstreeks bij de aanbieders verzamelen van de gegevens, dan zullen de aanbieders eveneens een pseudoniem moeten bijhouden dat is afgeleid van het BSN.

Net als in oplossingsrichting 3 is een knelpunt dat voor bestaande accounts geen BSN wordt bewaard door de aanbieders. Het opnieuw uitvragen van het BSN door de aanbieders (hierboven al genoemd als mogelijke oplossing) is (voor zover ons nu bekend) de enige oplossing om koppelbaarheid met RINpersoon te kunnen realiseren.

Relevante ontwikkelingen

Er spelen twee relevante ontwikkelingen die op de langere termijn andere mogelijkheden kunnen bieden binnen en tussen de beschreven oplossingsrichtingen:

- **De Europese *Data Governance Act*** regelt onder andere toegang tot data via databemiddelingsdiensten. Het CBS zou in deze context kunnen optreden als *Data Governing Body* en de datadeling kunnen faciliteren (in relatie tot de oplossingsrichtingen komt dat neer op een tussenvorm van richting 3 en 4). [Rij23]
- **De European Health Data Space Act** regelt het verbeteren van de beschikbaarheid van data in de zorgsector, via grensoverschrijdende infrastructuur. Naast patiëntgegevens direct ten behoeve van zorgverlening kunnen in dit kader ook andere gegevens (ook ten behoeve van preventie) worden uitgewisseld. [Jan25]

4.3 Afweging

Op de eerste oplossingsrichting na kunnen alle oplossingsrichtingen alle gestelde ontwerpdoelen invullen, zij het met ieder eigen mitsen en maren. Onderstaande Tabel 2 geeft een overzicht van de belangrijkste aandachtspunten.

Welke oplossingsrichting uiteindelijk de voorkeur heeft, is voor een groot deel een politieke keuze. Het is goed denkbaar dat specifieke elementen politiek relevant worden gevonden, en dat dit element de leidende factor wordt bij het maken van de afweging. De keuze voor een oplossingsrichting wordt dan ook niet in dit onderzoek gemaakt. Wel benoemen we hieronder een aantal eigenschappen van de verschillende oplossingsrichtingen die relevant kunnen zijn bij de afweging.

Doeltreffendheid en termijn

Binnen de kaders van dit onderzoek verwijst doeltreffendheid naar de mate waarin de gestelde ontwerpdoelen worden ingevuld. Het is uiteraard denkbaar dat er buiten de kaders van dit onderzoek *andere* doelen zijn waarvan wordt verwacht dat de uiteindelijke oplossing deze invult. Alle oplossingsrichtingen behalve de eerste komen in ieder geval tegemoet aan alle ontwerpdoelen, mits zij volledig worden uitgevoerd.

Omdat de oplossingsrichtingen bestaan uit deeloplossingen die gefaseerd in tijd kunnen worden uitgerold, bestaat de beleidskeuze met name uit de vraag in hoeverre de verschillende oplossingsrichtingen de doelen binnen de gewenste *termijn* treffen. Ons beeld hierbij is dat oplossingsrichtingen 2, 3 en 4 allen een behoorlijke tijdsinvestering vragen, om verschillende redenen: bij de een is tijd nodig voor het inrichten van

organisatorische processen, bij de ander is juist meer tijd nodig voor de totstandkoming van wet- en regelgeving. Bij richting 2 en 3 hangt de timing samen met die van wijziging van de Wok (die eventueel samenvallen met ongerelateerde wijzigingen).

Doelmatigheid

Met doelmatigheid wordt bedoeld op de efficiëntie van de oplossing: hoeveel maatschappelijke waarde ontstaat per geïnvesteerde euro? De eerste component (maatschappelijke waarde van het onderzoek dat wordt gefaciliteerd met de oplossingsrichtingen) is door ons niet becijferd, noch is dit in onze ogen goed mogelijk. De kosten voor het realiseren van de oplossingsrichtingen variëren eveneens sterk en zijn door ons niet in kaart gebracht.

Toch kunnen we op hoofdlijnen wel iets concluderen over de verschillen tussen de oplossingsrichtingen. Oplossingsrichting 1 (verbetering huidige werkwijze) is het 'lichtst' en kost relatief weinig geld, terwijl er twee belangrijke knelpunten worden opgelost. Het inrichten van een intermediair sluit (afhankelijk van de uitwerking) dicht aan op de huidige werkwijze en vraagt de minste aanpassingen bij aanbieders. De CBS-route en die op basis van een decentraal platform zijn het 'nieuwst', en kennen daarom (zo is onze verwachting) de hoogste eenmalige kosten. Kijken we naar de kosten per aanvraag, dan verwachten we dat deze bij de CBS-route het laagst zijn (de aanvrager betaalt het CBS hier overigens voor) en afhankelijk van de inrichting van een decentraal platform of de intermediair laag kunnen zijn (wanneer zo dicht mogelijk wordt aangesloten bij het CDB-datamodel).

Ruimte voor maatwerk

De oplossingsrichtingen verschillen in de mate waarin zij maatwerk faciliteren. In de eerste oplossingsrichting verandert dit niet ten opzichte van de huidige situatie. Bij een decentraal platform is alle maatwerk mogelijk, zolang de benodigde data via het platform wordt aangeboden (het platform faciliteert echter het aanbieden van ruwe data door aanbieders op een veilige wijze). Bij de routes intermediair en CBS is het datamodel meer rigide, wat betekent dat deze routes maatwerk in principe niet toestaan (het is wel denkbaar dat het datamodel periodiek wordt bijgesteld aan de behoefte). Het is de vraag of naast deze routes dan ook nog een maatwerkroute (in lijn met de huidige 31m-route) moet blijven bestaan.

Tabel 2 Overzicht van de mate waarin oplossingsrichtingen tegemoet komen aan de gestelde ontwerpdoelen (nadat de oplossingsrichting volledig is toegepast)

Ontwerpdoel	Oplossingsrichting 1 Verbeteren huidige bilaterale werkwijze	Oplossingsrichting 2 Decentraal platform	Oplossingsrichting 3 Centrale intermediair	Oplossingsrichting 4 CBS Microdata
✓ A Koppelen achtergrondkenmerken	✗ Niet mogelijk, omdat dit strijdig is met het vereiste (in artikel 31m Wok) dat moet worden geanonimiseerd.	✓ Mogelijk, maar koppeling met CBS (RINpersoon) vereist wel dat BSN opnieuw wordt uitgevraagd door aanbieders. Koppelen o.b.v. IBAN is wel mogelijk.	✓ Mogelijk, maar koppeling met CBS (RINpersoon) vereist wel dat BSN opnieuw wordt uitgevraagd door aanbieders. Koppelen o.b.v. IBAN is wel mogelijk.	✓ Mogelijk, maar koppeling met CBS (RINpersoon) vereist wel dat BSN opnieuw wordt uitgevraagd door aanbieders.
✓ D Verbreding onderzoeksdoelen	✗ Niet gerealiseerd, vanwege de beperking die is gesteld (en in deze richting blijft staan) in artikel 31m Wok.	✓ Gerealiseerd door wijziging artikel 31m Wok.	✓ Gerealiseerd door wijziging artikel 31m Wok.	✓ Gerealiseerd doordat het CBS zelf de gegevens opvraagt en beschikbaar stelt via Microdata.
✓ K Koppelen spelers tussen aanbieders	✗ Niet praktisch mogelijk; er kan bilateraal geen koppelbaar pseudoniem worden gehanteerd door de aanbieders.	✓ Mogelijk, maar vereist opnieuw uitvragen BSN door aanbieders, of het toewijzen van nieuwe onderzoekspseudoniemen.	✓ Mogelijk, maar vereist opnieuw uitvragen BSN door aanbieders, of het toewijzen van nieuwe onderzoekspseudoniemen.	✓ Mogelijk, maar vereist opnieuw uitvragen BSN door aanbieders, of het toewijzen van nieuwe onderzoekspseudoniemen.
✓ O Verbreding onderzoek door Ksa	✓ De Ksa kan na wijzigen van de taakstelling breder onderzoek uitvoeren.	✓ De Ksa kan na wijzigen van de taakstelling breder onderzoek uitvoeren.	✓ De Ksa kan na wijzigen van de taakstelling breder onderzoek uitvoeren.	✓ De Ksa kan na wijzigen van de taakstelling breder onderzoek uitvoeren.
✓ P Verbetering proces datatoegang	✓ Door de Ksa een actievere rol te laten invullen, ontstaat meer duidelijkheid bij beoordeling van aanvragen en iets meer afstand tussen onderzoeker en aanbieder.	✓ Het decentrale platform kan alle data-aanvragen en exports faciliteren. Een besispunt is in hoeverre aanbieders <i>zelf</i> goedkeuring geven, of dit door een centrale commissie gebeurt.	✓ De intermediair faciliteert het proces en beslist ook inhoudelijk over de aanvragen en het datamodel.	✓ Maakt gebruik van bestaande proces voor Microdata

Mate van openheid

De oplossingsrichtingen verschillen in de mate waarin deze open zijn voor onderzoekers en onderzoeken. Bij het decentrale platform kan op microniveau worden bepaald welke onderzoekers en onderzoeken welke data mogen gebruiken – het vraagstuk hier is met name wie deze keuze maakt (de aanbieders individueel of een centraal orgaan). In de intermediair-richting ligt deze keuze bij de intermediair en kan dit op een consistente en transparante manier plaatsvinden. In de CBS-route moet worden aangesloten bij de regels voor toegang tot CBS-microdata. In de praktijk verwachten we dat de beoogde doelgroep hier geen hinder van zal ondervinden.

Risico's

De oplossingsrichtingen verschillen ten aanzien van de risico's die er zijn dat de gedeelde data onverhoopt in verkeerde handen terecht komt. In de bilaterale werkwijze is het risico (zoals benoemd) met name dat de aanbieders van elkaar afwijkende inschattingen maken van onderzoeksaanvragen en bijbehorende afwegingen ten aanzien van dataminimalisatie en voorkomen van herleidbaarheid tot personen. Een foute inschatting leidt dan echter niet direct tot blootstelling van de gehele dataset, maar slechts het deel van de betreffende aanbieder.

Met moderne technologie kan in een decentrale oplossing een aantal soorten risico's op het gebied van cyberveiligheid worden uitgesloten, wat de oplossing veiliger kan maken dan andere oplossingen. Anderzijds zijn de technieken relatief nieuw en complex. Het is daarom cruciaal dat het systeem foutvrij en juist geconfigureerd is. Dit geldt overigens ook voor nieuw ingerichte centrale systemen.

Bij zowel de centrale intermediair als het CBS speelt dat vanwege de centralisatie van gegevens de impact van een gegevenslek in potentie het grootst is, omdat het meteen de gehele dataset kan treffen. Bij beiden zijn naast de veiligheid van de onderliggende technische systemen (voor o.a. de benodigde pseudonimisering) de organisatorische processen kritiek. Het splitsen van de taak van het toekennen van pseudoniemen van die van het uitwisselen van de gegevens biedt enige mitigatie. De CBS-richting heeft als voordeel dat die organisatie al lange tijd is ingericht op het verwerken en beschikbaar maken van dergelijke gegevens. Anderzijds kan worden beargumenteerd dat het vergroten van de hoeveelheid centraal beschikbare en koppelbare gegevens bij een overheidsorganisatie andere risico's met zich meebrengt.

5 Conclusie

5.1 Beantwoording onderzoeksvragen

Hoe ziet op hoofdlijnen de behoefte aan onafhankelijk onderzoek met maatschappelijk en/of wetenschappelijk doel op het gebied van speelgedrag bij kansspelen op afstand eruit, en wat is de daaruit voortvloeiende databehoeftte ten aanzien van beschikbare spelersdata?

Op basis van de verkenning onderscheiden we twee relevante categorieën van onderzoeken:

- **Gezondheidswetenschappelijke of *public health*-onderzoeken** richten zich op het in kaart brengen van gokgerelateerde problematiek op macroniveau. Deze onderzoeken monitoren en analyseren trends, risicogroepen en de prevalentie van probleemgedrag onder spelers. Dit is een fundamentele benadering om het ontstaan van speelgedrag en probleemgokken te begrijpen met of zonder de aanwezigheid van ander risicovol gedrag (bijv. alcoholgebruik, delinquentie).

De focus binnen dit type onderzoek ligt op geaggregeerde statistieken, waarbij het speelgedrag van grote groepen spelers wordt samengevat en geanalyseerd. Denk hierbij aan het totale aantal probleemspelers in een bepaalde periode, of het verloop van risicogedrag over tijd. Hoewel de benadering op macroniveau ligt, is er wel behoefte aan gedetailleerde uitsplitsingen. Relevant is om op te merken dat onderzoek in deze categorie in mindere mate *‘in de spellen zelf kijkt’*.

- **Gedrags- en neurologische onderzoeken** op microniveau richten zich op het analyseren van de impact van specifieke spelkenmerken of spelersgedragingen op kansspelerverslaving. Dit is een toegepaste benadering die zich richt op het beter begrijpen van verantwoord speelgedrag en de rol die specifieke spelontwerpen daarin spelen. Gedrags- en neurologische onderzoeken hanteren een ontwerpbenadering voor verantwoord gokken (of beleid) door een kader van gepersonaliseerde RG (Responsible Gaming) tools te ontwerpen (bijv. limieten, educatieve interventies, geplande pauzes, verschillende soorten en vormen van communicatie met spelers), om hun doeltreffendheid te vergroten door dergelijke maatregelen af te stemmen op de behoeften en voorkeuren van spelers. Voor de interpretatie van de data voor dit soort onderzoek is actieve medewerking van kansspel aanbieder essentieel: er wordt *‘in de spellen gekeken’* en in sommige gevallen zou zelfs een interventie in de spellen kunnen worden gedaan als onderdeel van een experiment.

Met name de eerste categorie onderzoeken is relevant voor de datadeling; in de tweede categorie is de databehoefte vrijwel altijd maatwerk, en is nauwe samenwerking met kansspelaanbieders onontkoombaar.

We merken tot slot op dat de behoefte aan data niet alleen kan voortvloeien uit onderzoeksbehoefte, maar beschikbaarheid van data kan ook leiden tot activering van latente onderzoeksbehoefte.

Kijken we naar de **soorten onderzoekers**, dan speelt met name een onderscheid op het aspect of onderzoekers ook *in opdracht van* de kansspelsector werken. In principe kan dit ook onderzoek betreffen dat relevant is voor beleidsvorming of de wetenschap. Het is ook niet ondenkbaar dat kansspelaanbieders in voorkomende gevallen een met de maatschappij c.q. wetenschap gedeeld belang hebben, bijvoorbeeld als het gaat om (onder andere) verslavingspreventie. De kansspelsector is immers ook gebaat bij het voorkomen van excessen en een positief imago. Het is ook aannemelijk dat dergelijk onderzoek in hoge mate onafhankelijk zou kunnen worden uitgevoerd, door het scherp definiëren van de rollen van en interactie tussen kansspelaanbieder en onderzoeker.

Uit de gesprekken volgt het beeld dat onderzoekers terughoudend zijn als het gaat om contact met de kansspelaanbieders. Buiten de genoemde risico's speelt hierbij de *perceptie* van niet-onafhankelijkheid een rol. Onafhankelijke onderzoekers willen ook de *schijn* voorkomen dat onderzoeken beïnvloed kunnen zijn door de kansspelsector.

Wat zijn op dit moment de juridische, technische en praktisch-organisatorische mogelijkheden tot het verstrekken van spelersdata om in de behoefte te voorzien?

In de Wet op de kansspelen (Wok [Wok]) worden vergunninghouders (in artikel 31m lid 5) verplicht om gegevens en analyses ten aanzien van speelgedrag *geanonimiseerd* beschikbaar te stellen voor onderzoek naar kansspelverslaving. Onderzoekers kunnen op grond van dit wetsartikel de aanbieders verzoeken om data.

Voor het opvragen van gegevens op basis van artikel 31m is door de Ksa in overleg met de sector een procedure ingericht. Hierbij krijgen onderzoekers twee maal per jaar de kans om een aanvraag in te dienen bij aanbieders. Voor het indienen van een aanvraag is een formulier opgesteld, dat beschikbaar is via de website van de Ksa.

Een aanvraag moet bilateraal worden ingediend (door een onderzoeker bij een specifieke aanbieder). De Ksa ontvangt daarnaast een afschrift van de verzoeken. Hoewel exacte gegevens ontbreken bij de Ksa, is het beeld op basis van gesprekken met de Ksa, aanbieders en onderzoekers, dat de hierboven beschreven procedure nog slechts zeer beperkt wordt gebruikt. Op moment van schrijven heeft één onderzoeker bij alle vergunde aanbieders een verzoek ingediend met betrekking tot de (volledige, geanonimiseerde) CDB-data.

Wat zijn op dit moment de juridische, technische en praktisch-organisatorische obstakels met betrekking tot het verstrekken van spelersdata om verder in de behoefte te voorzien?

In het onderzoek is een vijftal knelpunten geïdentificeerd, die maken dat de huidige wijze van datadeling ten behoeve van onderzoek niet aan de volledige (hier geïdentificeerde) databehoeftes vanuit onderzoek tegemoetkomt:

- **Het huidige proces voor datatoegang is suboptimaal ingericht.** Zowel onderzoekers als aanbieders zien het huidige proces als onpraktisch en kostbaar. Onderzoekers geven aan dat het feit dat slechts twee keer per jaar een aanvraag kan worden gedaan, onnodig beperkend is. Onhandig is daarnaast dat onderzoekers een aanvraag bij iedere aanbieder afzonderlijk moeten indienen. Er is daarnaast geen uniforme beoordeling van de verzoeken en geen uniforme controle op datakwaliteit.

Het huidige proces vereist daarnaast (te) nauwe bilaterale afstemming tussen onderzoeker en aanbieder. De relatie tussen onderzoekers en aanbieders is complex, waarbij enerzijds nauwe inhoudelijke afstemming nodig is, maar anderzijds een te nauwe samenwerking lastig is voor onderzoekers, vanwege de (schijn van) niet-onafhankelijkheid die het zou opleveren.

Vanuit het perspectief van de aanbieders is het belangrijkste bezwaar dat iedere aanvraag anders is en individueel moet worden behandeld, wat relatief veel werk en daarmee geld kost.

- **De gegevens over spelersgedrag zijn niet koppelbaar op spelersniveau tussen aanbieders.** Artikel 31m vereist dat aanbieders de verstrekte gegevens *anonymiseren*. Doordat aanbieders deze anonimisering individueel en onafhankelijk van elkaar doen is het per definitie niet mogelijk om gegevens over het speelgedrag van een speler bij de ene aanbieder te koppelen aan gegevens over speelgedrag van diezelfde speler bij een andere aanbieder. Hierdoor is het niet mogelijk om onderzoek uit te voeren waarin het gedrag van de speler *in samenhang* wordt bekeken. Dit is echter wel zeer relevant in het kader van (onder andere) onderzoek naar verslavingspreventie.
- **Er kunnen geen achtergrondkenmerken worden gekoppeld aan de data op spelersniveau.** Een ander gevolg van het vereiste om de op basis van artikel 31m gedeelde spelersgegevens te anonimiseren, is dat er ook geen andere datasets gekoppeld kunnen worden op spelersniveau. Voor sommige onderzoeken is het echter heel relevant om bepaalde achtergrondkenmerken van spelers te kunnen meenemen in een analyse.
- **De Ksa kan zelf niet alle onderzoeken uitvoeren die het zou willen.** De Ksa zou graag meer onderzoek willen doen. Het uitvoeren van onderzoek ten behoeve

van wetenschap of beleid is op dit moment géén wettelijke taak van de Ksa. De Ksa kan echter wel onderzoek uitvoeren ten behoeve van haar (bestaande) wettelijke taken. De wettelijke taakstelling (of althans de *interpretatie* daarvan door de Ksa) leidt echter tot een smallere afbakening van mogelijke onderzoeken dan gewenst. Het gebruik van de toezichtsdata (CDB) is op dit moment alleen mogelijk voor zover het onderzoek risicogestuurd toezicht betreft.

- **Toegang tot data via de huidige route is beperkt tot ‘onderzoek naar kansspelslaving’.** De verplichting uit artikel 31m heeft uitsluitend betrekking op onderzoek naar kansspelslaving. Daardoor zijn de gegevens niet altijd bruikbaar voor onderzoeken naar andere aspecten van de kansspelproblematiek, zoals het bredere begrip *gokschade*. Zowel vanuit de aanleiding voor dit onderzoek als vanuit de onderzoekers wordt duidelijk dat deze bredere vraag bestaat.

Om aanvullende gegevensdeling mogelijk te maken, waarmee voorgaande knelpunten worden geadresseerd, moeten een aantal juridische obstakels worden overkomen:

- **Rechtmatigheid van datadeling.** We hebben twee mogelijke obstakels hierbij geanalyseerd. Ten eerste speelt de vraag of de kansspelwetgeving zelf in de weg staat aan het delen van spelersgegevens op andere wijze dan via de bestaande verplichting van artikel 31m lid 5 Wok. Ten tweede speelt het verbod uit de AVG op het verwerken van bijzondere persoonsgegevens.

De kansspelwetgeving staat volgens ons niet in de weg aan het delen van niet-geanonimiseerde spelersgegevens. We kunnen dit niet met zekerheid stellen, omdat de regelgeving op dit punt niet geheel eenduidig is. De kansspelwetgeving bevat (art. 31m lid 5 Wok) een *verplichting voor aanbieders om mee te werken aan* een verzoek om geanonimiseerde gegevens ten behoeve van onderzoek naar kansspelslaving, maar verbiedt aanbieders volgens ons niet om niet-geanonimiseerde gegevens te verstrekken, noch om gegevens te verstrekken voor onderzoek naar een ander thema dan kansspelslaving.

Het gevolg is dat enkel nog het verbod uit de AVG op het verwerken van bijzondere persoonsgegevens resteert als juridisch obstakel voor het rechtmatig beschikbaar stellen van de gegevens. Hiermee kan worden omgegaan door ofwel (1) gebruik te maken van de uitzondering die de AVG op het verbod biedt ten behoeve van wetenschappelijk onderzoek, of (2) door de gegevens zodanig te bewerken dat geen sprake meer is van persoonsgegevens. Bij dat laatste merken we op dat koppelbaarheid over aanbieders ook gerealiseerd kan worden met geanonimiseerde gegevens. Door gebruik te maken van solide pseudonimisering, randomisering en generalisering, in combinatie met wettelijke waarborgen, kan namelijk worden gezorgd dat gegevens wel koppelbaar

zijn over aanbieders, maar niet meer redelijkerwijs herleidbaar zijn tot identificeerbare personen.

- **Medewerking aan de datadeling.** De aanbieders moeten worden bewogen tot medewerking, en/of (wanneer de aanbieders uit eigen beweging meewerken) moet deze medewerking worden geborgd. Het huidige wetsartikel borgt alleen de medewerking aan het huidige proces. Het is niet aannemelijk dat *alle* aanbieders uitsluitend op basis van vrijwilligheid meewerken aan een nieuwe datadeling. Door de kansspelregelgeving te wijzigen kunnen aanbieders worden verplicht om mee te werken aan de te kiezen oplossing. Een andere oplossing zou kunnen zijn om data door het CBS te laten verzamelen, die daarvoor over een speciale bevoegdheid beschikt.
- **Taakstelling Ksa.** Om mogelijk te maken dat de Ksa onderzoek doet dat verder gaat dan 'onderzoek ten behoeve van toezicht', moet de taakstelling van de Ksa in artikel 33b Wok worden aangepast. Zelfs als het takenpakket van de Ksa wordt aangepast, kan de controledatabank echter niet zomaar beschikbaar worden gesteld voor onderzoek. Daarvoor zou ook de regelgeving over de CDB zelf moeten worden aangepast.

Welke oplossingsrichtingen zijn er voor het beschikbaar stellen van spelersdata om verder (dan nu mogelijk is) in de onderzoeksbehoefte te kunnen voorzien?

De vijf bovengenoemde knelpunten kunnen worden vertaald naar ontwerpdoelen, en de juridische obstakels naar randvoorwaarden. Op basis hiervan zijn vier oplossingsrichtingen geformuleerd die aan deze ontwerpdoelen tegemoet komen en binnen de randvoorwaarden kunnen bestaan. Een oplossingsrichting zien we als een *geheel van technische, organisatorische en juridische deeloplossingen* dat (samen) zorgt dat tegemoet wordt gekomen aan (een deel van) de genoemde ontwerpdoelen.

In de praktijk bestaat er een sterke mate van afhankelijkheid tussen de onderliggend deeloplossingen: kies je voor een bepaalde technische invulling, dan is ook een bepaalde specifieke juridische oplossing nodig, en soms moet eerst een juridische stap worden gezet voordat een volgende stap kan plaatsvinden. Daarnaast hoeven niet alle ontwerpdoelen op hetzelfde moment in tijd te worden gerealiseerd. De beleidsmaker kan ervoor kiezen om een oplossingsrichting ten dele, of gefaseerd in tijd te effectueren. Het is bijvoorbeeld denkbaar dat op korte termijn de deeloplossingen ter verbetering van het proces worden ingezet, terwijl de koppelbaarheid met achtergrondgegevens iets is wat pas op langere termijn wordt gerealiseerd. Het kan echter wel zo zijn dat de keuze voor een bepaalde oplossingsrichting bepalend is voor hoe dat in de toekomst dan kan gebeuren.

De oplossingsrichtingen zijn de volgende:

1. **Verbeterde bilaterale werkwijze.** In deze oplossingsrichting wordt het huidige proces, waarbij een onderzoek bilateraal gegevens opvraagt bij aanbieders, verbeterd. De verbetering bestaat uit het toepassen van twee deeloplossingen: (1) het aanpassen van de taakstelling van de Ksa, zodat zij zelf meer vormen van onderzoek kan uitvoeren, en/of (2) het vergroten van de rol van de Ksa in het proces, zodat het proces soepeler kan verlopen.

Deze oplossingsrichting kan op relatief korte termijn worden gerealiseerd, en verbetert het huidige proces. Er wordt echter niet aan alle ontwerpdoelen tegemoet gekomen.

2. **Decentraal platform.** In deze oplossingsrichting wordt een nieuw platform ingericht waarop aanbieders data delen met onderzoekers. Dit gebeurt zoveel mogelijk decentraal. Met technische waarborgen wordt ervoor gezorgd dat partijen/personen die toegang hebben tot de data, en in principe beschikken over aanvullende informatie waarmee subjecten zouden kunnen worden geïdentificeerd, dit desondanks niet kunnen.

Met deze oplossing kan aan alle ontwerpdoelen tegemoet worden gekomen. Het technische platform is het meest flexibel, maar vraagt een behoorlijke eenmalige investering.

3. **Centrale intermediair.** Een derde partij (een nieuwe of bestaande organisatie, zoals de Ksa) zou kunnen optreden als verzamel- en uitgiftepunt voor de gegevens. Dit is in de basis een sterk organisatorisch gerichte oplossing: de derde partij beoordeelt onderzoekers en onderzoeksverzoeken, vertaalt deze naar dataverzoeken aan de aanbieders, en zorgt dat de gegevens uiteindelijk worden gekoppeld en geanonimiseerd ter beschikking komen van de onderzoeker. Binnen deze oplossingsrichting worden naast organisatorische/juridische instrumenten voor het borgen van de privacy ook technische instrumenten ingezet (primair pseudonimisering).

Deze oplossing vraagt ook een eenmalige investering, maar er kan in hogere mate worden aangesloten bij bestaande werkwijzen en systemen (bijvoorbeeld op het gebied van pseudonimisering). Hoewel wordt gewerkt met een generiek datamodel kan de oplossing voorzien in periodieke wijziging daarvan.

4. **CBS Microdata.** In deze oplossingsrichting worden de data verzameld door het CBS, en beschikbaar gemaakt in de CBS Microdata-omgeving. Onderzoekers kunnen onder strenge voorwaarden op projectbasis toegang krijgen tot deze omgeving om onderzoek uit te voeren. Het is mogelijk deze richting zo in te

richten dat de gegevens ook kunnen worden gekoppeld op persoonsniveau met andere datasets waarover het CBS beschikt, wat een belangrijk pluspunt kan zijn.

In deze richting wordt aangesloten bij de bestaande werkwijze en systemen van het CBS, maar moeten ten behoeve van koppelbaarheid en het uitwisselen van data tussen aanbieders en het CBS alsnog nieuwe processen worden ingericht. In deze oplossingsrichting bestaat de mogelijkheid om te koppelen met CBS-achtergronddata op persoonsniveau. Het datamodel is echter mogelijk meer rigide dan in de andere oplossingsrichtingen. Daarnaast moet worden aangesloten bij de toegangsvereisten van het CBS.

De afweging tussen de oplossingsrichtingen is aan de beleidsmaker. Daarbij speelt ons inziens primair de afweging tussen de kosten van de (deel)oplossingen en de (maatschappelijke) baten van het aanvullend mogelijk gemaakte onderzoek (beide aspecten zijn hier niet kwantitatief onderzocht). Daarnaast kunnen ook de realisatietermijn (o.a. ten gevolge van benodigde wijzigingen in wet- en regelgeving) en enkele andere aspecten (zoals de mate waarin wordt voorzien in maatwerk en de mate van openheid) en risico's ten aanzien van gegevensbescherming een rol spelen.

Verwijzingen

[AIVDRoseman]	AIVD (2024) <u>Deployment Advisory BSPA Assessment of Virtual Data Lake</u>
[APCovid]	Autoriteit Persoonsgegevens (2022) <u>Adviesverzoek onderzoek oversterfte. Brief aan de Tweede Kamer</u>
[APPseudoniemen]	Autoriteit Persoonsgegevens (17 maart 2025) <u>Gegevens pseudonimiseren</u>
[Baa04]	Baarsma, B. , Koopmans, C., Mulder, J., et al. (2004) <u>Goed(koop) geregeld: een kosten-batenanalyse van wetgeving en zelfregulering</u>
[Bes22]	<u>Besluit kansspelen op afstand</u> (2022)
[CBS22]	CBS (2022) <u>Richtlijnen voor Remote Access-output, augustus 2022</u> (geraadpleegd op 16 juli 2025)
[CBSMicrodata]	CBS <u>Catalogus microdata</u> (geraadpleegd op 16 juli 2025)
[CBSRA]	CBS (2021) <u>Aanvraag voor toegang tot microdata, versie 30 juli 2021</u>
[CBSRINPERSOON]	CBS (2023) <u>Documentatie Koppeltabel RINPERSOON met ABR voor zelfstandigen (KOPPELTABELZELFSTANDIGEN)</u> , CBS
[CIMS]	CBS (2 juli 2025) <u>Documentatie COVID-19 vaccinatiegegevens uit het COVID-vaccinatie Informatie- en Monitoringssysteem (CIMS)</u>
[Dia244]	Dialogic , Blom, T., Boiten, M., et al. (2024) <u>Drie jaar legaal online gokken. Evaluatie van de Wet Kansspelen op afstand (Koa)</u> , WODC
[Dia245]	Dialogic , te Velde, R., Hanswijk, M., et al. (2024) <u>Spelen met reclame. Het gebruik van promotionele kansspelen door vergunde kansspelaanbieders</u> , WODC
[EDPBPseudonymisation]	European Data Protection Board (EPDB) (2025) <u>Guidelines 01/2025 on Pseudonymisation. Adopted on 16 January 2025</u>
[ICTU22]	ICTU (2022) <u>BSN-koppelregister Polymorfe Pseudonimisering</u>
[Jan25]	Jansen, S. (27 maart 2025) <u>De EHDS is definitief: de zes belangrijkste implicaties</u> , ICTRecht
[Kan21]	Kansspelautoriteit (2021) <u>Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving</u>
[Kan22]	Kansspelautoriteit (17 maart 2022) <u>Verzoekformulier gegevens voor onderzoek naar kansspelverslaving</u> (geraadpleegd op 26 september 2025)
[Kan24]	Kansspelautoriteit (2024) <u>Onze strategie 2024 - 2028</u>
[Kan25]	Kansspelautoriteit (2025) <u>Monitoringsrapportage online kansspelen, voorjaar 2025</u>
[Kan251]	Kansspelautoriteit (2025) <u>Markers of Risk. How gambling data can be used to assist effective supervision.</u>
[Kir21]	Kirwan, M. , Mee, B., Clarke, N., et al. (2021) <u>What GDPR and the Health Research Regulations (HRRs) mean for Ireland: "explicit</u>

- [consent”—a legal analysis](#) in: Irish Journal of Medical Science (190) 515-521. doi:10.1007/s11845-020-02331-2
- [Kra24] **Kranenborg, H.R.** & Verhey, L.F.M. (2024) [De Algemene verordening gegevensbescherming in Europees & Nederlands perspectief](#), Wolters Kluwer
- [Ksa25] **Kansspelautoriteit** (2025) [Regeling verstrekking gegevens voor onderzoek naar kansspelverslaving](#) (geraadpleegd op 26 september 2025)
- [KVK] **KVK** [Rechtspersoonlijkheid](#)
- [Lin25] **Linksight** (2025) [Over ons](#)
- [Log24] **Logius** (10 januari 2024) [BSNk PP technische specificaties \(productie\)](#)
- [Min211] **Minister voor Rechtsbescherming** (26 januari 2021) [Besluit houdende bepalingen ter uitvoering van de Wet kansspelen op afstand \(Besluit kansspelen op afstand\)](#)
- [Min22] **Minister van Volksgezondheid, Welzijn en Sport** (18 november 2022) [Infectieziektenbestrijding. Brief aan de Tweede Kamer](#) in: Vergaderstukken Tweede Kamer, 2022-2023
- [Min233] **Minister voor Rechtsbescherming** (2023) [Kamerstuk. Brief van de minister van Rechtsbescherming.](#)
- [Pel22] **Pels Rijcken, Landsadvocaat** (2022) [Verstrekking van vaccinatiegegevens door het RIVM aan het CBS ten behoeve van statistisch onderzoek en terbeschikkingstelling van die gegevens aan academici ten behoeve van onderzoek naar de redenen en oorzaken van oversterfte](#)
- [Rij23] **Rijksoverheid** (29 september 2023) [Data Governance Act van kracht om data betrouwbaar en makkelijk te delen](#)
- [Ros] **Roseman Labs** [Technical Sheet](#)
- [SURF] **SURF** [SANE: veilige omgeving voor analyse van gevoelige data](#)
- [TNO PET] **TNO** (2023) [TNO - PET Lab - secure Multi-Party Computation \(MPC\)](#)
- [UAVG] [Uitvoeringswet Algemene verordening gegevensbescherming](#) (2021)
- [van09] **van Boom, W.H.**, Faure, M.G., Huls, N.J., et al. (2009) [Handelspraktijken, reclame en zelfregulering. Pilotstudy Maatschappelijke Reguleringsinstrumenten](#), WODC
- [van11] **van Boom, W.H.**, Lindenbergh, S.D., Philipsen, N.J., et al. (2011) [Energie-efficiency, afwikkeling van letselschade, algemene voorwaarden bij taxi en kinderopvang. Vervolgstudie maatschappelijke reguleringsinstrumenten](#), WODC
- [WCBS] [Wet op het Centraal bureau voor de statistiek](#) (2022)
- [Wok] [Wet op de kansspelen](#) (2024)
- [Wrvk] **Minister voor Rechtsbescherming** (2023) [Besluit werving, reclame en verslavingspreventie kansspelen](#)
- [Zwe25] **Zwenne, G.J.** (2025) [Commentaar op art. 4 AVG](#) in: T&C Privacy- en gegevensbeschermingsrecht, Wolters Kluwer

Bijlage 1. Overzicht interviewrespondenten en gesproken personen

Tabel 3 Overzicht interviewrespondenten

Organisatie	Functie
AmsterdamUMC	Associate professor
Autoriteit Persoonsgegevens ^a	Hoofd afdeling wetgeving en normuitleg Hoofd afdeling systeemtoezicht markt en bedrijfsleven
CBS	Innovatiemanager
Erasmus School of Social and Behavioural Sciences	Associate professor
IVZ (LADIS)	Programmanager
Jellinek	Manager preventie
Kansspelautoriteit (Ksa)	Product Owner CRUKS Programmamanager
Nederlandse Loterij	Compliance officer Managing director Adviseur public affairs Hoofd datamanagement
Roseman Labs	Oprichter
Trimbos-instituut	Lead researcher Digital Media & Gambling
Universiteit van Amsterdam (UvA)	Onderzoeker
VNLOK	Voorzitter Beleidsadviseur Beleidsmedewerker
ZonMw	Programmamanager
ZorgTTP	Functionaris gegevensbescherming Innovatiemanager

^a Zie voetnoot a (pagina 7)

Tabel 4 *Overzicht deelnemers validatiesessie*

Organisatie	Functie
CBS	Innovatiemanager
Kansspelautoriteit (Ksa)	Research Coordinator Adviseur strategie
Ministerie van Justitie en Veiligheid	Beleidsmedewerker Integriteit en Kans- spelen
Nederlandse Loterij	Hoofd datamanagement
Universiteit van Amsterdam	Onderzoeker

Bijlage 2. Casuïstiek

Om inzicht te krijgen in hoe andere domeinen omgaan met vergelijkbare vraagstukken rondom het beschikbaar stellen van op persoonsniveau koppelbare gegevens voor onafhankelijk onderzoek, zijn enkele relevante casussen onderzocht. Daarbij is specifiek gekeken naar de gekozen aanpak, de knelpunten die daarbij zijn ondervonden en de wijze waarop deze zijn ondervangen. In de onderstaande figuur is de werkwijze schematisch weergegeven.



Inventarisatie van casussen

Voorafgaand aan de selectie is gekeken naar casussen in verschillende sectoren. Een aantal relevante casussen die uit deze inventarisatie naar voren kwamen zijn:

- Het LADIS (Landelijk Alcohol en Drugs Informatie Systeem) is een door het IVZ samengestelde dataset met data afkomstig van instellingen voor verslavingszorg. De dataset wordt in opdracht van het ministerie van VWS samengesteld en beheerd door IVZ.
- De Nederlandse Zorgautoriteit (NZa) verzamelt zorgdeclaraties van diverse zorgaanbieders, onder meer ten behoeve van de doorontwikkeling van zorgprestatie modellen. Vektis is een 'shared business intelligence center' van de zorgverzekeraars. Vektis verzamelt declaratiedata van zorgverzekeraars en deelt deze data met verschillende zorginstanties, waaronder de NZa.
- In de openbaarvervoersector worden reisgegevens verzameld door Translink, de organisatie achter de OV-chipkaart. Deze data wordt onder voorwaarden beschikbaar gesteld voor onderzoeksdoeleinden. Daarnaast loopt er in de provincies Groningen en Drenthe een pilot vanuit Publiek Vervoer waarbij reisgegevens binnen het regionale OV onder concessie eveneens worden verzameld en gedeeld voor analyse en beleidsondersteuning.
- Het Centraal Bureau voor de Statistiek (CBS) verzamelt ruwe gegevens op persoonsniveau in het kader van haar wettelijke taak. Deze data bestrijkt verschillende maatschappelijke domeinen, zoals arbeid, gezondheid, onderwijs en inkomen. Geaccrediteerde onderzoekers kunnen binnen een beveiligde omgeving analyses uitvoeren op deze data.

Selectie van casussen

Voor de selectie van relevante casussen is een aantal beoordelingscriteria opgesteld die relevant zijn voor het voorliggende vraagstuk. Aan de hand van dit kader is

beoordeeld in hoeverre een casus naar verwachting bruikbare lessen kan bieden. Specifiek is gezocht naar aanwezigheid van de volgende elementen in de casus:

- Of er sprake is van **koppelbaarheid** van data op persoonsniveau in data afkomstig van verschillende dataleveranciers. Indien sprake is van dergelijke koppelbaarheid, biedt de case aanknopingspunten voor hoe wordt omgegaan met (her)pseudonimisering en met het risico op depseudonimisering, met name door de dataleveranciers zelf.
- Of er **bijzondere persoonsgegevens** worden verwerkt. Daaraan worden van wetswege strenge eisen gesteld. Dergelijke casussen kunnen waardevolle inzichten bieden in hoe deze eisen in de praktijk worden toegepast en georganiseerd.
- Omdat **onafhankelijkheid** is benoemd als aandachtspunt in de verkenning van de onderzoeksbehoefte, is per case onderzocht in hoeverre en op welke manier de onafhankelijkheid van onderzoek wordt geborgd door middel van organisatorische oplossingen. Relevante casussen beperken het gebruik van de beschikbare data tot specifieke doelen, zoals maatschappelijk of wetenschappelijk onderzoek. Hierbij is onderzocht hoe en in welke mate deze **gebruiksdoelen** van de data **beperkt** worden.
- Tot slot is er gekeken of data-aanbieders **verplicht zijn om data aan te leveren**, ongeacht of deze verplichting wettelijk is vastgelegd of op een andere manier is georganiseerd.

Figuur 5 toont het selectiekader voor de case studies, afgezet tegen de vooraf opgestelde beoordelingscriteria. Op basis hiervan zijn LADIS, CBS Microdata en Publiek Vervoer Groningen Drenthe geselecteerd als casussen voor nadere uitwerking. Een uitgebreidere toelichting op de geselecteerde casussen volgt onder de figuur.

	In welke mate is de data op persoonsniveau koppelbaar over verschillende data-leveranciers?	Is er sprake van bijzondere persoonsgegevens?	In welke mate wordt de onafhankelijkheid van het onderzoek gewaarborgd?	Is het verplicht voor "aanbieders" om data aan te leveren?	In welke mate worden de case gebruiksdoelen van de data beperkt (tot bijvoorbeeld maatschappelijk en wetenschappelijk onderzoek)?
Beoogd voor gokdata	Volledig	Ja	Hoog	Ja	Hoog
LADIS	Ja	Ja	Hoog	Ja	Ja
OV Translink	Niet volledig	Nee	Middel	Ja	Hoog
NZA	Niet	Ja	Hoog	Ja	Hoog
CBS-microdata	Volledig	Ja	Hoog	Ja	Hoog
Vektis	Niet volledig	Ja	Laag	Nee	Hoog
Publiek vervoer Groningen Drenthe	Volledig	Ja	Hoog	Ja	Hoog

Figuur 5 Selectiekader en invulling daarvan voor de verschillende gevonden casussen

Geselecteerde casussen

CBS microdata

Het CBS ontvangt in veel gevallen gegevens op basis van het burgerservicenummer (BSN), waarna deze intern worden gepseudonimiseerd. Hierdoor kunnen datasets aan elkaar worden gekoppeld zonder dat direct herleidbare persoonsgegevens worden gebruikt. Het CBS verwerkt verschillende soorten bijzondere persoonsgegevens, waaronder gezondheidsinformatie en gegevens over etnische afkomst. Onderzoekers krijgen alleen toegang tot microdata wanneer het CBS, onafhankelijk van dataleveranciers, heeft vastgesteld dat het betreffende onderzoek voldoet aan de voorwaarden. De dataleveranciers hebben geen inspraak in deze beoordeling. De aanlevering van gegevens aan het CBS is wettelijk verplicht op grond van de Wet op het Centraal Bureau voor de Statistiek.^a Toegang tot de data wordt uitsluitend verleend voor onderzoek dat kwalificeert als wetenschappelijk of statistisch.

LADIS

LADIS is een publieke organisatie en heeft als doel gegevens ten behoeve van de verslavingszorg te verzamelen in opdracht van het ministerie van VWS. De koppelbaarheid van gegevens wordt mogelijk gemaakt doordat een pseudoniem wordt aangemaakt op basis van het BSN, dat mag worden verwerkt op grond van een wettelijke basis. Het betreft bijzondere persoonsgegevens, aangezien het gaat om inschrijvingen bij een verslavingsinstelling en daarmee om gezondheidsgegevens. Stichting IVZ beoordeelt onafhankelijk of een onderzoeker een bepaald onderzoek mag uitvoeren;

^a [\[wetten.overheid.nl\]](http://wetten.overheid.nl)

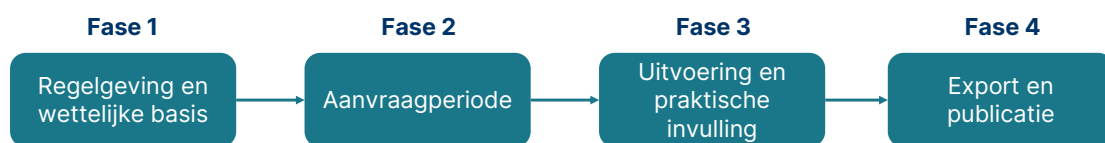
dataleveranciers hebben hierin geen inspraak. De gegevens worden gedeeld met een derde partij, ZorgTTP. De aanlevering van gegevens is verplicht op grond van de Wet kwaliteit, klachten en geschillen zorg (Wkkgz). Het gebruik van de data is beperkt tot specifiek omschreven doelen, zoals monitoring, beleidsontwikkeling en wetenschappelijk onderzoek binnen het domein van de verslavingszorg.

Publiek Vervoer Groningen Drenthe

De gegevens van het reisgedrag vanuit de Wet maatschappelijke ondersteuning (Wmo) van verschillende vervoerders binnen de pilot Publiek Vervoer Groningen Drenthe (PVGd) zijn koppelbaar op basis van klantnummers. Omdat het onderzoek zich richt op onder andere vergrijzing in de regio's en Wmo-rittenbanken, waar gegevens indicaties van gezondheid of specifieke doelgroepen kunnen bevatten, gaat het om bijzondere persoonsgegevens en bepaalde achtergrondkenmerken van personen. De gegevens worden gedeeld met een derde partij Roseman Labs. De aanlevering van deze gegevens is verplicht vanuit de vervoersovereenkomst die met vervoerders is afgesloten. De data wordt uitsluitend gebruikt voor publieke doelen, zoals het toegankelijk houden van het openbaar vervoer en het inspelen op toekomstige ontwikkelingen zoals vergrijzing.

Analyse van geselecteerde casussen

Voor de analyse van de casussen is een analysekader als uitgangspunt genomen. Het analysekader is opgebouwd rond een viertal opeenvolgende fases die zijn te onderscheiden in een onderzoeksproces op basis van extern verkregen data. In elke fase is gekeken naar hoe randvoorwaarden, knelpunten en oplossingen zijn ingericht of toegepast in de geselecteerde casussen. Het analysekader is in onderstaande figuur weergegeven.



In fase 1 is onderzocht welke grondslag van toepassing is op de verwerking van (bijzondere) persoonsgegevens, en of er sprake is van een wettelijke verplichting tot aanlevering van data. Fase 2 richt zich op de procedure voor het aanvragen van toegang tot data. Hierbij wordt gekeken naar de eisen die worden gesteld aan onderzoekers of onderzoeksinstellingen, de criteria waarop aanvragen worden beoordeeld (zoals maatschappelijke doeleinden), wie deze beoordeling uitvoert, en de doorlooptijd van aanvraag tot goedkeuring. Fase 3 betreft de uitvoering van het onderzoek met de beschikbare data. In deze fase is geanalyseerd op welke wijze de pseudonimisering van persoonsgegevens is ingericht, of de data centraal wordt opgeslagen en of onderzoekers werken binnen een afgeschermd omgeving. Daarnaast is gekeken of er kosten verbonden zijn aan het gebruik van de data en hoe lang de

doorlooptijd is tussen het indienen van een exportaanvraag en het daadwerkelijk verkrijgen van de export. In fase 4 is onderzocht of de exportdata wordt gecontroleerd door de data-verstrekker, bijvoorbeeld op dataminimalisatie of identificeerbaarheid. Verder is gekeken of er een verplichting geldt tot publicatie van de onderzoeksresultaten.

Fase 1. Regelgeving en wettelijke basis

Bij CBS Microdata zijn dataleveranciers op grond van een AMvB verplicht gegevens aan te leveren. CBS is bevoegd om ten behoeve van statistische doeleinden gebruik te maken van deze gegevens volgens artikel 33 van de Wet op het Centraal Bureau voor de Statistiek (Wet CBS). Voor de verwerking van persoonsgegevens door het CBS is geen toestemming van individuele betrokkenen vereist. De grondslag hiervoor is de uitvoering van de publieke taak (Artikel 6(1)(e) AVG en Wet CBS).

Ten behoeve van LADIS is er eveneens een wettelijke verplichting voor zorgaanbieders voor de aanlevering van gegevens op grond van de Wet kwaliteit, klachten en geschillen zorg (Wkkgz). Voor de verwerking van persoonsgegevens door LADIS vindt plaats op basis van een wettelijke grondslag, waardoor er geen toestemming van individuele gebruikers is vereist.

Bij PVGD zijn de vervoerders op basis van de vervoersovereenkomst verplicht medewerking te verlenen. De reiziger stemt door de aanschaf van de vervoerspas in met het delen van data voor onderzoeksdoeleinden.

Fase 2. Aanvraagperiode

Het CBS beoordeelt op een groot aantal criteria zowel de onderzoeker of onderzoeksinstelling (voor het verkrijgen van een instellingsmachtiging) als de aanvraag. Het CBS voert onder andere een privacytoets uit, waarbij onder meer wordt beoordeeld of er sprake is van een geldige AVG-grondslag en een duidelijke doelbinding voor statistische of wetenschappelijke doeleinden. Het streven is om binnen één maand na indiening van de aanvraag toegang tot de data te verlenen.

Bij LADIS is op dit moment geen sprake van externe toegang voor onderzoekers tot data, vanwege recente wijzigingen ten aanzien van de wettelijke grondslag. Voorheen bestond er wel een grondslag die toegang mogelijk maakte voor onderzoekers. Er wordt gekeken naar het opnieuw openen van deze route. In het verleden werd zowel de onderzoeker c.q. onderzoeksinstelling als de data-aanvraag beoordeeld door experts van IVZ. IVZ verkent of deze toetsing in de toekomst door externe beoordelaars zou kunnen worden uitgevoerd. De aanvraag wordt getoetst op basis van een aantal vaste criteria. Voorafgaand aan de levering van de data wordt een gebruikersovereenkomst opgesteld, waarin heldere afspraken worden vastgelegd over het doel van het datagebruik, de minimale omvang van data, een zo hoog mogelijk aggregatieniveau, het voorkomen van (ook indirecte) herleidbaarheid tot individuen, het uitsluitend

gebruik van geaggregeerde rapportages en een beperkte bewaartermijn. Buiten deze afspraken vindt geen aanvullende privacybeoordeling plaats. Afhankelijk van de complexiteit van de aanvraag varieert de doorlooptijd van 2 tot 6 weken.

Bij de pilot PVGD werd de aanvraag beoordeeld door functionarissen gegevensbescherming van de deelnemende gemeenten. Zij voeren ook de privacybeoordeling uit van de partij die de data verstrekt. De screening van Roseman Labs, de partij die de data voor PVGD verwerkt, is uitgevoerd door het Ministerie van Infrastructuur en Waterstaat. Omdat de Wmo-ritgegevens in beheer zijn van de gemeenten, kunnen verschillende partijen deze data opvragen, mits dit gebeurt binnen de geldende privacyregels en voor legitieme doeleinden zoals onderzoek of beleidsontwikkeling.

Fase 3 Uitvoering en praktische invulling

Het CBS vervangt direct identificerende gegevens, zoals naam, adres en Burgerservicenummer (BSN), zo snel mogelijk na ontvangst door een unieke en verder betekenisloze code, het zogenaamde RIN-nummer. Het CBS heeft aan iedere persoon in de GBA (Gemeentelijke Basisadministratie) een RIN toegekend (zie paragraaf 4.2.4). Op basis van een BSN-nummer of op basis van andere unieke kenmerken kan het CBS een persoon in het GBA opzoeken en de bijbehorende RIN bepalen. [CBSRINPERSOON]

De sleutel om deze RIN's te vertalen naar en van de oorspronkelijke BSN/RSIN wordt veilig bewaard bij het CBS en is niet toegankelijk voor onderzoekers. Het RSIN is het Rechtspersonen en Samenwerkingsverbanden Informatienummer. Hierdoor zijn de gegevens waarin personen met een RIN worden geïdentificeerd niet direct herleidbaar tot individuele personen. Omdat de CBS-datasets echter zeer gedetailleerd en uitgebreid zijn, is indirecte herleidbaarheid (op basis van de beschreven kenmerken in de dataset) onverminderd mogelijk.

Alle microdata wordt centraal opgeslagen en beheerd binnen de beveiligde CBS Microdataomgeving. Onderzoekers werken binnen een afgeschermd projectomgeving waarin de benodigde datasets beschikbaar zijn. Alle gegevens die buiten de CBS Microdataomgeving gebracht worden, worden als 'export' beschouwd. Voordat gegevens geëxporteerd mogen worden, worden ze door het CBS beoordeeld op onthullingsrisico. De doorlooptijd hiervoor is doorgaans 3 werkdagen.

LADIS maakt gebruik van de diensten van ZorgTTP voor de verwerking van persoonsgegevens. Deze gegevens worden lokaal (bij de data-aanbieder) omgezet naar pseudoniemen via een Privacy Verzend Module (PVM). De gebruikte pseudoniemen zijn voor ZorgTTP ontsleutelbaar. De gepseudonimiseerde gegevens worden vervolgens verzonden naar ZorgTTP, waar de pseudoniemen een tweede keer (en nu onomkeerbaar) worden versleuteld. De gegevens die aanbieders aanleveren zijn echter niet uniform. Een inhoudscommissie is een adviesorgaan dat de inhoud van kwaliteitsregistraties beoordeelt, zich buigt over de vraag hoe items beter geoperationaliseerd

kunnen worden en legt dit vast in specificaties. Hierdoor wordt de datakwaliteit en vergelijkbaarheid verbeterd. De inhoudscommissie is breed samengesteld met vertegenwoordigers van instellingen en andere betrokkenen.

Er bestaat de mogelijkheid voor onderzoekers om met de data te werken binnen een afgeschermd omgeving. In de wet (Wkkgz) is geregeld dat LADIS de databron vormt voor een aantal partijen, waaronder EUDA, het Trimbos-instituut en het CBS. Voor aanvragen ten behoeve van onderzoek worden de kosten van de volledige procesinzet, van aanvraag t/m einde onderzoek (vernietiging data), in rekening gebracht. Afhankelijk van de complexiteit van de vraagstelling en of het een herhaalde aanvraag betreft, varieert de doorlooptijd van de export tussen 1 en 4 weken.

Bij PVGD wordt gebruikgemaakt van Secure Multi-Party Computation (SMPC), waarmee koppelingen tussen verschillende databronnen kunnen worden gemaakt op basis van klantnummers. SMPC is een techniek waarmee meerdere partijen gezamenlijk berekeningen kunnen uitvoeren op hun data, zonder dat zij elkaars data hoeven te delen of in te zien. De gegevens blijven gedurende het hele proces versleuteld en zijn niet toegankelijk voor de andere partijen. Alleen de vooraf overeengekomen partij(en) krijgen inzage in de uitkomsten van de berekening. De SMPC-oplossing van PVGD wordt als platform (dienst) geleverd door Roseman Labs. Om herleidbaarheid tot individuele personen te voorkomen, geldt als voorwaarde dat resultaten alleen worden opgeleverd wanneer deze gebaseerd zijn op een minimale 'bucket' (celvulling) van 10 records. De data wordt centraal beheerd door PVGD. Onderzoekers werken binnen een afgeschermd omgeving, zodat de data niet op hun eigen systemen beschikbaar is. Er is bij PVGD geen sprake van een doorlooptijd tussen het indienen van een exportaanvraag en het verkrijgen van de export.

Fase 4 Export en publicatie

Onderzoekers zijn verplicht om bij het CBS een ingevuld outputformulier toe te voegen aan hun exportaanvraag. Het CBS beoordeelt vervolgens binnen drie werkdagen of de aangevraagde output voldoet aan de privacyrichtlijnen. Daarnaast geldt bij het CBS een verplichting tot publicatie van onderzoekresultaten op basis van CBS Microdata. Onderzoekers dienen binnen vier weken na publicatie de referenties van alle publicaties aan het CBS te verstrekken.

Bij LADIS wordt eveneens de exportdata gecontroleerd op herleidbaarheid tot individuele personen. Er geldt echter geen verplichting tot publicatie van onderzoeksresultaten. Ook bij PVGD bestaat geen verplichting om onderzoeksresultaten te publiceren. De dataverstrekker stelt bepaalde eisen aan de exportdata, maar er vindt geen inhoudelijke beoordeling van de exportdata plaats.



Onderzoek voor *onderbouwd* beleid.

Dialogic innovatie & interactie

Hooghiemstraplein 33

3514 AX Utrecht

030-2150580

www.dialogic.nl