

Vergadering Afnemersraad MiDaS
Maandag 14 april 2014 14.30-16.30 uur
Locatie: Panteia, Zoetermeer

Aanwezig:

Paul de Graaf (OBI Rotterdam)
Dirk Jan Griffioen (RIVM)
Bas van der Klaauw (VU)
MARIKE Knoef (Netspar)
Jan van der Linden (Panteia)
Johan Polder (RIVM)
Ineke Stoop (SCP, voorzitter)
Ilonca Vaartjes (UMC Utrecht, Julius Centrum)
Bas ter Weel (CPB)

Ruben Dood (hoofd Centrum voor Beleidsstatistiek)
Ruurd Schoonhoven (Relatiebeheerder MiDaS, secretaris)

Verhinderd:

Michiel van den Hauten (Min. SZW)
Bob van Waveren (Regioplan)

Agenda:

1. Opening, vaststellen agenda, welkom nieuwe leden
(*Verslag vergadering 7 oktober 2013 is al vastgesteld en op website geplaatst*)
2. Invoering nieuwe richtlijnen van de CCS voor toegang tot microdata
3. Controle remote access werkplekken
4. Onderzoek op microdata in internationaal perspectief
5. Gang van zaken rond outputcontroles
6. Evaluatie tarieven en tariefstructuur
7. Actiepunten vorige vergadering
8. Overige mededelingen
9. Rondvraag en sluiting

1. Opening, vaststellen agenda, welkom nieuwe leden

De voorzitter heet als nieuwe leden welkom Dirk Jan Griffioen (vanaf vandaag opvolger van Johan Polder als representant van RIVM) en Marike Knoef (representant van Netspar). Notulen van de vorige vergadering zijn reeds per mail becommentarieerd en goedgekeurd en inmiddels op de website van het Centrum voor Beleidsstatistiek gepubliceerd. Een aanvulling op de agenda ingebracht door Michiel van den Hauten wordt behandeld na punt 8

2. Invoering nieuwe richtlijnen van de CCS voor toegang tot microdata

In de achterliggende periode is de reeds in de vorige vergadering van de Afnemersraad besproken revisie van de richtlijnen voor toegang tot CBS microdata afgerond. De belangrijkste criteria blijven dat microdata alleen gebruikt mogen worden voor statistisch of wetenschappelijk onderzoek en dat de resultaten van dat onderzoek voor de openbaarheid gepubliceerd moeten worden. Er wordt onderscheid gemaakt tussen een toets op de instelling en de toets op het onderzoeksproject. Voor onderzoeksprojecten in opdracht van een betalende derde opdrachtgever dient altijd toestemming van de DG te worden verkregen.

Ten aanzien van de instellingen geldt dat nu ook de in de CBS-wet met name genoemde instellingen als planbureaus, universiteiten en andere bij wet ingestelde onderzoeksinstellingen, aan wie de DG op eigen gezag toegang kan verlenen, eenmalig getoetst dienen te worden. Voor overige (veelal private) instellingen dient een machtiging te worden aangevraagd van de CCS, en deze instellingsmachtiging kan nu ook voor een periode van 5 jaar worden afgegeven.

De nieuwe regeling is per 1 maart j.l. van kracht geworden. De omschrijving van criteria en procedure is op internet gepubliceerd, met de nu definitieve vragenlijsten voor de instellings- en projecttoets. Ten aanzien van de punten waarvoor de afnemersraad aandacht heeft gevraagd (zie notulen van het vorige overleg):

- Bedoeling is dat de administratieve lasten voor afnemers beperkt blijven door instellings- en projectvragenlijsten geheel te scheiden.
- Er wordt nog een web-vragenlijst of vergelijkbaar ontwikkeld.
- Voor reeds gemachtigde instellingen zal de nieuwe procedure naar verwachting niet tot extra doorlooptijd leiden, behalve voor opdrachtonderzoeken indien die aanleiding geven tot nadere vragen; waakzaamheid is op dit punt geboden.
- Criteria voor toetsing van projecten: vooralsnog niet meer dan de gebruikelijke CBS-kaders. Bedoeling is om dit te verhelderen aan de hand van op te bouwen 'jurisprudentie'.
- Publicatieplicht: zie onder.
- Lopende 5-jaars machtigingen vervallen, alle instelling moeten opnieuw een machtiging aanvragen. Er worden daarbij geen problemen verwacht voor reeds gemachtigde instellingen waar geen sprake is van veranderingen.
- De secretaris van de CCS zal een registratie van nieuwe aanvragen bijhouden.
- Communicatie: alle actieve gebruikers, projectleiders et cetera zijn per e-mail geïnformeerd over de nieuwe procedure. Een aantal grote afnemers is daarnaast gericht benaderd.

In de discussie vraagt de afnemersraad opnieuw aandacht voor doorlooptijden bij offertes voor opdrachtonderzoek, die zou niet langer dan een week moeten zijn. Vervolgens komt de vraag aan de orde of een derde-geldstroomproject (dat wil zeggen gefinancierd door een collectebusfonds of vergelijkbaar) bij een universiteit wordt gezien als een 'eigen' onderzoek dan wel een onderzoek in opdracht. Ruben Dood neemt dit op met de CCS. Ook wordt gesproken over de eisen rond publicatie. Uiteraard hoeft niet ieder detail gepubliceerd te worden, maar wel op zijn minst de hoofdlijnen en belangrijkste conclusies, onderbouwd met de betreffende cijfers. Zijn er geen publicabele resultaten, dan graag een bericht aan CvB.

Veel wetenschappelijk tijdschriften verlangen dat ook de data waarop de publicatie is gepubliceerd worden meegeleverd, een eis waaraan in dit geval niet voldaan kan worden. De data blijven wel binnen het CBS beschikbaar mocht heranalyse of replicatie van het onderzoek nodig zijn. Er zijn voorbeelden van andere statistische bureaus voor formuleringen hieromtrent die in een publicatie kunnen worden opgenomen. Ruben Dood zal dit in samenwerking met Ilonca Vaartjes uitwerken en een daartoe strekkende zinsnede zal vervolgens in de documentatierapporten worden opgenomen in de al bestaande paragraaf over bronvermelding.

3. Controle remote access werkplekken

Ruben Dood bericht dat de CBS-directie en CCS verlangen dat remote access werkplekken periodiek worden bezocht om te controleren of zij aan de eisen voldoen. Instellingen met laag risico (dat wil zeggen waar toegang tot het gebouw gecontroleerd wordt middels persoonsidentificatie) zullen eens per 5 jaar bezocht worden, instellingen met vrije publieke toegang (universiteiten, UMC's) eens per jaar. Bij buitenlandse instelling worden foto's van de remote access werkplek opgevraagd.

4. Onderzoek op microdata in internationaal perspectief

Ruben Dood geeft een korte presentatie van verschillende initiatieven. Het project DARA van Eurostat, dat streefde naar een centrale repository van microdata van verschillende landen, is niet tot ontwikkeling gekomen doordat de technische voorzieningen en daarmee de financiering noch de accreditatie goed genoeg geregeld kon worden, onder andere omdat de meeste landen hun microdata hiervoor niet willen/kunnen afstaan.

Het EU programma Data Without Boundaries (DWB) is in zijn eerste fase wat CBS betreft succesvol in de zin dat momenteel een tiental remote access onderzoeksprojecten vanuit diverse landen gebruik maakt van CBS microdata. In andere landen is dit minder van de grond gekomen.

CBS werkt voorts aan een pilot met enkele Scandinavische landen om 'bottom-up' tot een werkbaar model te komen, uitgaande van wat in de betrokken landen intern al werkt, en met oog voor de aspecten die in eerder genoemde initiatieven onvoldoende aan bod komen. Het vervolgetraject hiervan,

samen met een Europese partner zoeken naar een technische en administratieve structuur voor 'secure multiparty computation' lijkt veelbelovend.

De vergadering ondersteunt het idee om te proberen vanuit een kleine werkende situatie verder op te schalen, in plaats van meteen groots te beginnen. Er wordt aandacht gevraagd voor de mogelijkheid om, naast economisch onderzoek, ook bijvoorbeeld medisch onderzoek op internationale schaal mogelijk te maken; dit kwam in DWB nauwelijks aan bod. Ruben Dood vermeldt nog dat de ervaring tot nu toe is dat er met de bestaande mogelijkheden wel vergelijkende studies tussen landen worden gedaan, maar dat concrete voorbeelden van echt pan-Europees onderzoek (waarin data van verschillende landen in één gemeenschappelijke studie worden betrokken) nog ontbreken. Marike Knoef noemt nog LISSY, een systeem rond de Luxembourg Income Study waarbij data van verschillende landen via remote execution toegankelijk zijn binnen de vertrouwelijkheidsregels van die landen.

5. Gang van zaken rond outputcontroles

Ruben Dood geeft een korte presentatie, en begint met het kader aangegeven door art. 37 van de CBS-wet, dat stelt dat '.....gegevens slechts zodanig openbaar worden gemaakt dat daaraan geen herkenbare gegevens over een afzonderlijk persoon, huishouden, onderneming of instelling kunnen worden ontleend'. De gegevensbescherming rond microdata onderzoek betreft aan de voorkant van een project de CCS machtiging, geheimhoudingsverklaring, toetsing van onderzoekers, doelbinding en strikte login procedures. Na uitvoering van het onderzoek volgt outputcontrole. Dit alles maakt dat de restricties in de onderzoeksomgeving heel beperkt kunnen blijven. In de praktijk wordt 80% van de aangeboden outputs direct geaccepteerd, 15% wordt vrijgegeven na enkele kleine aanpassingen die door CBS-outputcontroleurs, veelal in overleg met de onderzoekers, kunnen worden aangebracht. De resterende 5% wordt tegengehouden. Juist deze groep is, door overleg intern-CBS en discussie met onderzoekers, dikwijls zeer tijdrovend. Redenen voor tegenhouden van outputs kunnen meer of minder zwaarwegend zijn (onthulling door kleine aantallen in tabellen of scatterplots, als output aanbieden van RIN of BEID nummers(!), groepsonthulling of dominantie, in strijd met de richtlijnen zoals gepubliceerd op de website). Niettemin zijn er ook voorbeelden van argumenten waarom soms tabellen met kleine aantallen toch niet onthullend zijn.

De discussie gaat over de vraag hoe gebruikers de outputcontroles ervaren en hoe zij daarmee omgaan.

Hoewel outputcontroles in de praktijk vooral als lastig worden ervaren komt uit de raad ook positief geluid ten aanzien van de uitleg in gevallen waarin outputs niet (zonder meer) kunnen worden doorgelaten. Ook wordt opgemerkt dat strenge outputcontroles als gunstig neveneffect hebben dat zij onderzoekers dwingen goed na te denken over de analyses die ze willen doen. Sommige onderzoekers sparen outputs op om budgettaire redenen, andere maken bewust gebruik van meekijkers zodat een onderzoeker samen met zijn begeleider op de RA werkplek tussenresultaten kan bespreken zonder de kosten voor outputcontrole.

Er blijkt belangstelling te zijn voor een snelle en goedkope controle voor kleine tabellen die aan nader te omschrijven voorwaarden voldoen, en geen gegevens bevatten uit speciaal gevoelige bestanden zoals rond doodsoorzaken, criminaliteit en dergelijke. Ruben Dood zal een voorstel uitwerken.

6. Evaluatie tarieven en tariefstructuur

Nadat gebleken was dat de per 1-1-2012 ingevoerde kostenstructuur en bijbehorende tarieven had geleid tot een fors tekort over het jaar 2012 zijn de tarieven per 1-1-2013 bijgesteld, binnen dezelfde structuur. Inmiddels is gebleken dat CvB over het jaar 2013 hiermee kostendekkend heeft kunnen opereren.

Niettemin wil CvB alert zijn op mogelijke behoeften aan aanpassingen van de tariefstructuur bij gebruikers daar waar onevenwichtigheden worden ervaren. Het gaat dan bijvoorbeeld om:

- Door toevoeging van verschillende nieuwe diensten is het aantal 'knoppen waaraan onderzoekers kunnen draaien' om de kosten van hun project te sturen verder toegenomen, maar dit is de transparantie van het geheel niet ten goede gekomen.

- De harde grenzen van de projectklassen leiden ertoe dat onderzoekers nogal eens concessies moeten doen met betrekking tot het aantal bestanden, omdat de enkele extra bestanden die ze eigenlijk nodig hebben leiden tot onevenredig veel hogere kosten.

Sommige diensten worden door de leden van de raad als onevenredig kostbaar ervaren, zoals het toevoegen van nieuwe jaargangen van een bestand waarvan het project de eerdere al in gebruik heeft. Uit de raad komt de suggestie om een vast bedrag voor een dataset (bestandsonderwerp) te hanteren, en dan nieuwe jaargangen kosteloos toe te voegen, met startbedrag eventueel gedifferentieerd naar eenvoudige of complexe bestandsonderwerpen. Het feit dat bestanden niet van een project verwijderd mogen worden, ook al zijn ze bij het begin van het project toegevoegd 'voor de zekerheid' maar zijn ze nooit gebruikt, wordt als erg rigide ervaren.

Ten aanzien van zeer grote projecten wordt opgemerkt dat er ten onrechte geen rem meer is op het toevoegen van nog een onderzoeker of nog meer bestanden; maar anderzijds dat het opsplitsen van een groot project niet altijd leidt tot bijvoorbeeld drie kleine projecten, maar dikwijls tot drie eveneens grote projecten; en dan is niet duidelijk welk doel met opsplitsen gediend is.

Tot slot wordt van verschillende kanten gevraagd om duidelijker bijlagen bij de facturen.

CvB neemt de reacties van de raad mee; dit onderwerp komt bij de volgende vergadering terug.

7. Actiepunten vorige vergadering

Punten die nog niet aan de orde zijn geweest:

- Informatie over nieuwe bestanden en bestandsonderwerpen: op de website (bestandencatalogus) staan sinds enige tijd aparte categorieën 'recent verschenen' en 'nieuwe onderwerpen'. Daarnaast wordt informatie over nieuw verschenen bestandsonderwerpen opgenomen in de elektronische nieuwsbrief die enkele malen per jaar aan alle onderzoekers wordt toegezonden.
- Er zijn enkele meedenkers/testers voor een WIKI aangemeld, maar de vorderingen op dit punt gaan traag, doordat ICT aspecten rond licenties en dergelijke veel tijd hebben gekost.

8. Overige mededelingen

- Binnen het CBS wordt een discussie gevoerd over de zorgvuldigheidseisen die moeten worden gesteld indien een onderzoeker een eigen microdatabestand in zijn project wil laten importeren om dat te koppelen aan CBS data. In het projectcontract is hierom de volgende zinsnede opgenomen: *"De onderzoeker verklaart dat de gegevens rechtmatig zijn verkregen en dat ook de verstrekking van de gegevens aan het CBS voldoet aan de eisen van de Wet bescherming persoonsgegevens en, voor zover het gaat om gezondheidsgegevens, (mede) aan de eisen van afdeling 7:5 van Boek 7 van het Burgerlijk Wetboek (Wet geneeskundige behandelingsovereenkomst, WGBO). Het CBS kan hierover nadere informatie opvragen."* Voorkomende gevallen worden voorgelegd aan de CBS jurist.
- Eveneens wordt binnen het CBS gesproken over de vraag in hoeverre CvB mede verantwoordelijk kan zijn voor cq kan toezien op de kwaliteit van de uitkomsten van microdata onderzoek. Conclusie is en blijft dat wij output alleen controleren op onthullingsrisico, en dat het aan de onderzoekers is om bij geringe aantallen waarnemingen zich af te vragen wat de resulterende cijfers waard zijn. CvB verwacht verder van onderzoekers dat zij zich houden aan de 'do's and dont's' genoemd in de documentatierapporten, die dikwijls ook over kwaliteit en bruikbaarheid van de data gaan.
- Op 8 april j.l. is de 100^e remote access computer geïnstalleerd.

(toegevoegd agendapunt)

Michiel van den Hauten (SZW) heeft geïnformeerd naar welke software CBS op de langere termijn gaat gebruiken, want op dit moment is de software van CBS niet compatibel met de software van het Rijk. Dit leidt tot problemen voor RA werkplekken bij SZW. Kern van het probleem is dat het Rijk als conventie (DWR) hanteert Windows 7 64 bits, terwijl CBS-RA werkt in een omgeving van Windows 7 met 32 bits applicatie. Aangezien SZW in een DWR omgeving zit is afwijken van het gebruikelijke besturingssysteem niet mogelijk.

Ruurd Schoonhoven licht toe dat het zeker geen beleid van het CBS is om aan 32 bits systemen vast te houden, maar dat we te maken hebben met de ook voor CvB zeer onbevredigende situatie dat de betreffende leveranciers tot nu toe niet in staat of bereid zijn gebleken 64-bits versies van drivers van de vingerafdruklezers en van de bijbehorende biometrische authenticatie software te leveren. Hiervoor is nog geen definitieve goede oplossing gevonden; helaas is de enige en hopelijk tijdelijke oplossing om hiervoor een 32-bits omgeving onder Windows 7 64 bits te emuleren. Er wordt gekeken naar andere vormen van biometrische identificatie, ook naar andere vingerafdruklezers, maar een dergelijke keuze zou tot gevolg hebben dat alle honderden gebruikers opnieuw geregistreerd moeten worden; dit is daarom ook geen aantrekkelijke optie.

8. Rondvraag en sluiting

- Paul de Graaf maakt melding van het congres van de de VSO (contactorgaan van gemeentelijke statistisch onderzoekers) op 24 april a.s. waar ook aandacht gegeven zal worden aan de diensten van het Centrum voor Beleidsstatistiek, zowel met betrekking tot maatwerk- als microdata-onderzoek.
- Voorzitter en hoofd CvB danken scheidend lid Johan Polder voor zijn inbreng in de discussies en zijn inzet voor de afnemersraad over de afgelopen drie jaar.

Actiepunten

- Ruben Dood neemt de status van projecten gefinancierd door derde geldstroom op met de CCS.
- Ruben Dood werkt in samenwerking met Ilonca Vaartjes een formulering in relatie tot de bronvermelding uit die aangeeft dat data bij het CBS beschikbaar blijven voor replicatie van het onderzoek, maar niet aan bijvoorbeeld een wetenschappelijk tijdschrift ter beschikking kunnen worden gesteld.
- Ruben Dood zal een voorstel uitwerken voor snelle, goedkope outputcontroles.
- In de volgende afnemersraad zal de tariefstructuur en de kostenstructuur voor bestanden op de agenda staan. CBS bereidt hiervoor een stuk voor.
- Ruben Dood gaat na of verduidelijking van de factuurbijdragen mogelijk is.