



Notitie voor rondetafelgesprek Vaste commissie voor Infrastructuur en Waterstaat over Toegang tot OV-data d.d. 10 september 2019

Het CBS voorziet de samenleving van relevante en betrouwbare statistische informatie. Daarmee wordt inzicht in maatschappelijke vraagstukken geboden en wordt het debat, de beleidsontwikkeling en besluitvorming gevoed. In het debat in de Tweede Kamer over de beschikbaarheid van ov-data is gevraagd of het CBS een rol kan spelen bij het ontsluiten van informatie over ov-reizigersstromen. Deze notitie gaat in op de rol en taken van het CBS, de beschikbare data-infrastructuur, de ontsluiting van databronnen en het beschikbaar stellen van open data. Daarbij wordt nader toegelicht hoe het CBS gegevensbescherming en privacy borgt. Vervolgens wordt ingegaan op de mogelijkheden van ov-data als belangrijke potentiële bron voor statistieken. Om te kunnen voorzien in de informatiebehoefte van de samenleving wil het CBS toegang krijgen tot de ov-data.

Wettelijke taak CBS: onafhankelijke statistieken voor praktijk, beleid en wetenschap

Overall in de samenleving groeit de vraag naar betrouwbare cijfers voor onderbouwing en verantwoording van besluitvorming. Burgers hebben behoefte aan transparante, feitelijke informatie. Overheid en bedrijven zijn op zoek naar actuele, nauwkeurige en gedetailleerde informatie zodat zij hun beleid en bestuur kunnen onderbouwen met betrouwbare cijfers. Wetenschappers gebruiken data als basis voor hun onderzoek. Het CBS heeft een wettelijke taak om van overheidswege onafhankelijke statistieken te produceren en publiceren voor praktijk, beleid en wetenschap. Met betrouwbare en samenhangende statistische informatie biedt het CBS inzicht in maatschappelijke vraagstukken.

Dataknooppunt van de Nederlandse overheid

Om betrouwbare statistieken te kunnen maken en tegelijk de administratieve lasten voor burgers en bedrijven te verlagen, heeft het CBS bij wet toegang gekregen tot overheidsdata. Het CBS heeft de beschikking over meer dan 200 dataregisters van meer dan 100 overheidsinstellingen. Zodoende worden data die al door andere overheidsorganisaties zijn vergaard, hergebruikt, verrijkt en gecombineerd voor statistische doeleinden. Het CBS is hierdoor een dataknooppunt van de Nederlandse overheid geworden waar alle soorten databronnen kunnen worden gecombineerd om te worden verwerkt tot bruikbare statistische informatie die bijvoorbeeld ingezet kan worden voor beleidsvraagstukken in overheidsbelang.

Open data

Alle door het CBS geproduceerde statistieken zijn beschikbaar in de vorm van interactief te raadplegen tabellen via de databank StatLine en ook als machine leesbare "open data". Open data zijn vrij toegankelijke datasets die eenvoudig door computers verwerkt kunnen worden. Het gaat om data die voor iedereen toegankelijk en herbruikbaar zijn zonder beperkingen. Via het CBS data portaal zijn ruim 4500 datasets beschikbaar. Geautomatiseerd hergebruik van de uitkomsten wordt door deze open-data-structuur mogelijk, zowel voor het CBS zelf als voor externe partijen. Vanuit open data kunnen veel nieuwe soorten diensten en apps ontwikkeld worden voor allerlei toepassingen.



Gegevensbescherming en privacy

Het beveiligen van gegevens en het beschermen van de privacy van personen en bedrijven is één van de kernwaarden van het CBS. Waarborgen voor privacy en gegevensbescherming zijn opgenomen in de CBS-wet. Het CBS heeft uitgebreide protocollen en procedures ingericht om privacy en gegevensbescherming te borgen. Het CBS laat ook geregeld externe audits op allerlei aspecten van informatiebeveiliging uitvoeren en het CBS blijft voortdurend tegen het licht van nieuwe technologische ontwikkelingen toetsen of de privacy is geborgd.

Ontsluiten van open data voor andere overheidsinstanties

Het CBS is één van de weinige partijen in Nederland die data van andere partijen ontvangt en deze data, al dan niet gecombineerd, beschikbaar maakt als open data, via een combinatie van kanalen. Een belangrijke stap daarbij is dat privacygevoelige data door middel van statistische methoden en technieken worden omgezet in veilige informatie. Dit is noodzakelijk om te voorkomen dat informatie naar individuele personen of bedrijven herleidbaar is. Het CBS kan andere overheidspartijen hierdoor ondersteunen bij hun opendata-ambities. Bij het CBS is veel kennis van en ervaring met databeveiliging, privacybescherming en het publiceren van open data.

Mogelijkheden voor gebruik van microdata voor onderzoek

In toenemende mate hebben praktijk, beleid en wetenschap behoefte aan de brongegevens van het CBS die ten grondslag liggen aan de statistieken. Deze brongegevens, de zogenoemde microdata, zijn combineerbare data op persoons-, bedrijfs- en adresniveau. Deze data zijn bijzonder privacygevoelig en om die reden zijn bij microdata-onderzoek strikte voorwaarden en strenge beveiligingseisen van toepassing. Uitsluitend daartoe geautoriseerde instellingen kunnen onder strikte voorwaarden onderzoek doen op de microdatabestanden van het CBS. Deze gegevens worden gepseudonimiseerd beschikbaar gesteld en de data verlaten nooit de beveiligde IT-omgeving van het CBS. Een van de vereisten is dat de onderzoeksresultaten worden gepubliceerd, zodat ook die informatie beschikbaar wordt gesteld voor de samenleving.

Voorzien in nieuwe informatiebehoefte samen met het CBS

Om de steeds complexer wordende samenleving adequaat te kunnen blijven beschrijven is het CBS steeds op zoek om het dataknooppunt met nieuwe informatiebronnen uit te breiden. Hierbij zoekt het CBS nadrukkelijk samenwerking met publieke en private partijen¹. Technologische en maatschappelijke veranderingen leiden tot nieuwe informatiebehoefte. Tegelijkertijd neemt door digitalisering de hoeveelheid beschikbare data explosief toe. Denk hierbij aan data beschikbaar van bijvoorbeeld sociale media, sensoren in wegen, uit smartphones en van het internet. Door innovatieve methoden kunnen nieuwe (big) databronnen worden gebruikt en kan de juiste informatie, in de juiste vorm, bij de juiste persoon op het juiste moment terecht komen. Daarbij gaat het steeds meer om het beschrijven van fenomenen in de samenleving aan de hand van met elkaar samenhangende indicatoren.

¹ Daarbij hoeven overigens niet alle data bij het CBS aanwezig te zijn. Met nieuwe technieken op het gebied van “veilig data delen”, zoals data virtualisatie-technologie, is de locatie van de data minder belangrijk geworden.



Bij de ontwikkeling van nieuwe dataproducten ziet het CBS grote voordelen in samenwerking met het bedrijfsleven en kennisinstellingen. De samenwerking in onderzoek en innovatie wordt vormgegeven in een data-ecosysteem. Daarin wordt de data-infrastructuur en kennis van het CBS gecombineerd met de innovatiekracht, specialistische kennis en mogelijkheden van private partijen en academische instellingen. De samenwerking met Translink is hier een voorbeeld van.

OV-data belangrijke bron voor statistieken

De ov-chipkaart levert een schat aan reisdata op. Deze data kunnen een essentiële aanvullende bron zijn voor de CBS-statistieken over personenmobiliteit in Nederland voor nu en in de toekomst. Toegang tot deze ov-data maakt het mogelijk om bestaande CBS-statistieken² over het vervoer per spoor naar het gewenste kwalitatieve niveau te brengen en via het dataknooppunt medeoverheden te voorzien in hun informatiebehoefte aangaande mobiliteit in een veranderende samenleving. Het CBS heeft zoals vermeld bij wet voor statistische doeleinden de beschikking over alle registraties van de overheid. De ov-data vallen daar echter niet onder. Om te kunnen voorzien in de informatiebehoefte van de samenleving wil het CBS dus toegang krijgen tot deze ov-data.

Waardevolle informatie en inzichten

Het CBS kan de ov-data koppelen aan andere bij het CBS beschikbare databronnen, zoals gegevens over de plaats waar mensen wonen, werken en onderwijs volgen of bijvoorbeeld gegevens over autobezit of rijbewijzen. Dat kan voor overheden en vervoerders nog meer waardevolle inzichten opleveren in het gebruik van het openbaar vervoer, reispatronen en reismotieven van bepaalde groepen in de Nederlandse samenleving (denk aan forenzen, studenten en ouderen). Gedetailleerde ov-gegevens bieden ook mogelijkheden om regionale uitsplitsingen te maken voor reizigerskilometers naar provinciaal en gemeentelijk niveau. Zo kan per regio worden bekeken hoeveel gebruik wordt gemaakt van het openbaar vervoer in relatie tot de behoeften van de reizigers. De nieuwe statistische informatie kan een basis bieden voor monitoring en evaluatie van beleidsmaatregelen, zoals maatregelen met betrekking tot infrastructurele voorzieningen of projecten in het kader van Mobility as a Service. De koppeling van databronnen biedt ook mogelijkheden om onderzoek te doen naar de keuzemogelijkheden en het gedrag van de reizigers.

Samenwerking CBS met Translink en ontwikkeling landelijke ov-monitor

Het CBS is met Translink in gesprek over het beschikbaar stellen van ov-chipkaartdata. Daarbij kijken het CBS en Translink hoe op een veilige en verantwoorde manier informatie kan worden samengesteld uit deze data. Hiervoor doen CBS en Translink gezamenlijk onderzoek naar de

² Veel CBS-statistieken zijn gebaseerd op Europese verordeningen die de lidstaten verplichten om gegevens te leveren aan Eurostat, het statistiekbureau van de Europese Unie. Een van deze statistieken betreft het vervoer van goederen en personen per spoor. Een belangrijk onderdeel van de in 2018 herziene Europese Spoorverordening vormt de informatie over "reizigersvervoer per spoor". Op basis van deze verordening en op grond van zijn wettelijke taak, kan het CBS deze informatie bij de betrokken vervoerders uitvragen. Echter, in het kader van het minimaliseren van de lasten voor bedrijven is het efficiënter als het CBS de gewenste informatie via Translink ter beschikking kan krijgen.



mogelijke toepassingen van ov-chipkaartdata in combinatie met andere databronnen om nieuwe statistische indicatoren te ontwikkelen of bestaande indicatoren te verbeteren.

Momenteel wordt in een pilot samengewerkt met Translink aan de ontwikkeling van een landelijke ov-monitor, waarbij ov-chipkaartinformatie wordt gecombineerd met gegevens van het CBS. Vervoerders, overheden en gebruikers worden nauw betrokken bij de samenstelling van de ov-monitor en de uitkomsten worden als open data gepubliceerd. Met de gebruikers van deze monitor zal bekeken worden of het continueren of uitbreiden van de monitor een helder overheidsbelang dient en of de betrokken overheden de kosten daarvan structureel kunnen vergoeden.