

aan DG

cc Directeur CSB, CPO, hoofd CAD, privacycoördinatoren

van FG, plvFG

onderwerp Dataminimalisatie

datum 26 juni 2023

Managementsamenvatting

Dataminimalisatie is een van de beginselen die ex art. 5 AVG van toepassing zijn op de verwerking van persoonsgegevens. Het implementeren van dataminimalisatie is een verplichting onder de AVG en vereist een bewuste inspanning om elke verwerking van persoonsgegevens, waaronder veel van de statistische processen bij het CBS, zo veel mogelijk te beperken. In dit advies is uitgewerkt welke dimensies van dataminimalisatie in processen bij het CBS kunnen worden onderscheiden. Het advies dient ter ondersteuning van beleidsvorming en als hulpmiddel bij de evaluatie of vormgeving van processen.

Aanleiding

Dit advies vloeit voort uit mijn toezichtsagenda over 2022. Dataminimalisatie is een van de beginselen waaraan verwerkingsverantwoordelijken zich conform de AVG dienen te houden.

Wettelijk kader

Het beginsel van dataminimalisatie is vastgelegd in artikel 5 lid 1 onder c: “[Persoonsgegevens moeten] toereikend zijn, ter zake dienend en beperkt tot wat noodzakelijk is voor de doeleinden waarvoor zij worden verwerkt („minimale gegevensverwerking”);”. Overweging 156 geeft extra context voor de interpretatie van het beginsel in statistische verwerkingen. Deze stelt dat technische en organisatorische maatregelen moeten worden genomen om passende waarborgen voor de rechten en vrijheden van betrokkenen te bieden bij onder andere verwerkingen met statistische doeleinden. De tekst van art. 5 lid 1 onder c betekent onder andere dat als een doel bereikt kan worden zonder de verwerking van persoonsgegevens, die verwerking niet is toegestaan. Derhalve is dataminimalisatie geen ‘nice to have’ maar een ‘need to have’ onder de AVG.

In artikel 89 lid 1 AVG wordt pseudonimisering genoemd als een van de technische en organisatorische maatregelen die kunnen worden getroffen om het beginsel van minimale gegevensverwerking te garanderen, onder de voorwaarde dat het statistische onderzoek¹ hiermee nog steeds kan plaatsvinden. In de laatste zin van het artikellid gaat de AVG nog een stapje verder: als het doel van de verwerking kan worden bereikt terwijl identificatie van betrokkenen niet langer

¹ Het artikel betreft ook andere doeleinden: archivering, wetenschappelijk of historisch onderzoek. Ik laat deze verder onbesproken.



mogelijk is, moet de verwerking op die wijze plaatsvinden. Dataminimalisatie wordt dan een afpelprincipe:

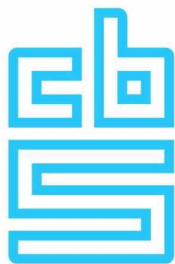
- Indien een statistisch doel kan worden bereikt waarbij de (her)identificatie van betrokkenen niet langer nodig is, dan moet worden geanonimiseerd
- Indien een statistisch doel kan worden bereikt met verwerking van gepseudonimiseerde gegevens, dan is pseudonimisering verplicht
- Indien een statistisch doel alleen kan worden bereikt door verwerking van niet-geanonimiseerde gegevens, dan zijn anonimisering noch pseudonimisering verplicht.

In het bovenstaande mag het beveiligen van data door deze te anonimiseren niet verward worden met het pseudonimiseren van data. Omdat bij pseudonimisering gegevens gekoppeld kunnen worden met unieke personen is de AVG van toepassing. Dit geldt niet voor geanonimiseerde gegevens, omdat deze per definitie niet meer herleidbaar zijn tot personen. Het anonimiseringsproces valt als verwerkingshandeling wel onder de AVG. Dit betekent dat er een verplichting rust op het CBS om pseudonimisering en anonimisering (het laatste ook, maar niet alleen, in de vorm van aggregatie en outputcontrole) effectief toe te passen. Dit betekent eveneens, dat het CBS moet kijken naar alternatieven voor het huidige regime van verrinning, omdat door het lange en brede gebruik van dit 'pseudoniem' de effectiviteit niet meer gegarandeerd kan worden.

Nota bene: in de verschillende stappen van een statistisch proces kan een verschillend regime van toepassing zijn. Zo zal bij enquêtes doorgaans het individu kenbaar moeten zijn om de verzamelde gegevens later te kunnen koppelen aan achtergrondkenmerken. Zo kan bijvoorbeeld in een enkel proces eerst met persoonsgegevens, vervolgens met gepseudonimiseerde persoonsgegevens, en vervolgens met anonieme gegevens worden gewerkt.

Dataminimalisatie is verbonden aan het beginsel van Privacy by Design, dat is neergelegd in art. 25 AVG. Dit artikel stelt in lid 1 dat gegevensbeschermingsbeginselen, zoals dataminimalisatie, moeten worden nagestreefd met behulp van technische en organisatorische maatregelen. Deze verplichting geldt zowel bij de bepaling van de middelen van de verwerking als bij de verwerking zelf. De verwerkingsverantwoordelijke, die voor deze maatregelen verantwoordelijk is, bepaalt deze met inachtneming van de stand van de techniek, de uitvoeringskosten, aard, omvang, context en het doel van de verwerking. In de te maken overweging worden ook de waarschijnlijkheid en ernst van de risico's voor betrokkenen meegenomen.

Dataminimalisatie raakt eveneens het proportionaliteits- en subsidiariteitsbeginsel. Het proportionaliteitsvereiste brengt met zich dat het doel van de verwerking van de persoonsgegevens in verhouding moet staan tot de inbreuk op de privacy van de betrokkene. Dit betekent bijvoorbeeld ook dat persoonsgegevens niet langer mogen worden bewaard dan noodzakelijk. In dat kader is het volgen van het bewaartermijnenbeleid belangrijk. Het CBS conformeert zich daarin aan de AVG en aan de Archiefwet. Het subsidiariteitsbeginsel gaat over de vraag of het doel waarvoor de persoonsgegevens worden verwerkt niet op een andere, voor de betrokkenen minder nadelige, wijze kan worden verwekelijkt.



Het CBS beschikt over zo veel persoonsgegevens dat het risiconiveau vrijwel per definitie hoog tot zeer hoog is, en daarmee ook de lat voor de te treffen maatregelen hoog ligt. De risico's vloeien onder meer voort uit het zeer veelomvattende beeld dat in potentie van een betrokkene kan worden gevormd, en de aard van de gegevens (bijzonder, strafrechtelijk, ID-nummers). Dit vergt de voortdurende aandacht voor technische en organisatorische maatregelen om de risico's te mitigeren. Ook betekent dit dat wat twintig jaar geleden nog 'state of the art' was, dat vandaag niet per definitie meer is.

Guidance

In haar richtlijnen uit 2020² geeft de European Data Protection Board (EDPB) nadere uitleg over de interpretatie van art. 25 van de AVG over data protection by design and by default. Omdat art. 25 ziet op het treffen van technische en organisatorische maatregelen die uitvoering moeten geven aan onder andere het beginsel van dataminimalisatie, zijn deze richtlijnen ook relevant voor dit FG-advies. De meest relevante punten zijn de volgende:³

- Hoeveelheid verzamelde persoonsgegevens: betreft het volume, de soorten, categorieën en het detailniveau (granulariteit)
- De mate waarin zij worden verwerkt: beperking tot noodzakelijke verwerking qua soort en frequentie
- De termijn waarvoor zij worden opgeslagen: opslag alleen in geval dit noodzakelijk is in relatie tot het doel van de verwerking
- De toegankelijkheid daarvan: beperking van toegang tot de personen en de soorten toegang die noodzakelijk zijn

Governance

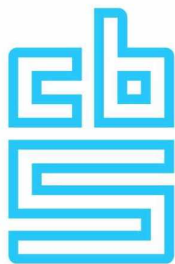
Het beginsel van dataminimalisatie behoeft controle in de CBS-organisatie. De vanuit de AVG opgelegde beperkingen dienen immers door de organisatie heen gehandhaafd te worden. Door het beginsel te vertalen in beleid, en het beleid zo veel mogelijk meetbaar te operationaliseren per divisie of zelfs per sector, ontstaat accountability of uitlegbaarheid van de uitvoering van de AVG. Zowel de naleving als de aantooningsplicht zijn vastgelegd in art. 5 lid 2 van de AVG.

Dimensies van dataminimalisatie in het statistisch proces

Zoals vrijwel alle regels uit de AVG is het beginsel van dataminimalisatie een zogenaamde 'open norm'. Zo'n norm houdt ook onder technologische en maatschappelijke ontwikkelingen stand. De interpretatie ervan verschuift met de veranderingen die plaatsvinden. De volgende opsomming van dimensies van dataminimalisatie is dan ook niet uitputtend bedoeld, maar als een gids bij het nagaan van verbeteringsmogelijkheden in (onder andere) statistische processen als de dataverzameling, dataverwerking en voorbereiding voor publicatie.

² Guidelines 4/2019 on Article 25 Data Protection by Design and by default. Version 2.0. Adopted on 20 October 2020.

³ Idem, p. 12-14.



Dimensie	Minimalisatie bij verzamelen en verwerken van persoons- en bedrijfsgegevens door
Tijd	Verkorten leverings- en verwerkingsperiode Verlagen periodiciteit levering Beperken toegang in de tijd Beperken bewaartermijn
Variabelen	Alleen variabelen verzamelen/gebruiken die nodig zijn voor het statistische doel Minder bijzondere persoonsgegevens verzamelen/gebruiken Aggregeren, anonimiseren, (effectief) pseudonimiseren van direct identificerende gegevens (waar mogelijk en in die volgorde) Grootteklassen/categorieën van variabelen gebruiken Beschikbare variabelen bij gegevensverzameling hergebruiken Statistisch beveiligen van output
Opslag	Verminderen aantal rustpuntbestanden/versies Verwijderen niet-noodzakelijke duplicaten op verschillende locaties Beperken/uitbannen fysieke verstrekking
Toegang	Beperken wie, waar, hoe en hoe lang toegang krijgt tot persoonsgegevens (lees-/wijzigrechten, verdere verwerkingsmogelijkheden) Compartimenteren toegang gebruikers (van verschillende functiegroepen) in variabelen en records

Aanbevelingen

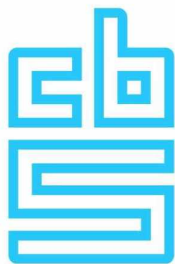
Op grond van het bovenstaande kom ik tot de volgende aanbevelingen, waarvan het management binnen de geschetste kaders zelf de prioritering en implementatie kan bepalen:

Governance

- Evalueer en herzie periodiek het beleid en de implementatie ervan op het gebied van dataminimalisatie.
- Zorg voor adequaat beleid en implementatie daarvan voor logging en monitoring van gegevenstoegang. Herzie periodiek de inrichting van processen met inachtneming van de relevante beginselen zoals dataminimalisatie.
- Pas het beginsel van dataminimalisatie eveneens toe binnen de bedrijfsstatistieken en bij de facilitaire processen van het CBS.
- Zorg voor adequate controle en allocatie van verantwoordelijkheden bij de verdere verstrekking van gegevens binnen en buiten het CBS en via RA.

Centralisatie

- Centrale bestanden en voorzieningen als het ABR, SSB en DSC maken het eenvoudiger om dataminimalisatiebeleid te handhaven; bespoedig het gebruik van deze voorzieningen en ban tegelijkertijd het fenomeen van 'schaduwadministraties' (eigen kopieën) van microdatabestanden uit.



- Zorg voor voldoende mitigering van risico's die (juist) voortvloeien uit centralisatie door middel van compartimentering en andere technische en organisatorische maatregelen.

Eenvoud

- Beperk het aantal processtappen zo veel mogelijk, automatiseer de processtappen om menselijke fouten zo veel mogelijk te kunnen voorkomen
- Reduceer het aantal verschillende omgevingen waarin gewerkt kan worden en geef duidelijke richtlijnen voor elk van deze omgevingen
- Beperk voor zover mogelijk het aantal uitzonderingen op staand beleid, bijv. op het gebied van softwaregebruik, omgevingsgebruik, BSN-gebruik etc.
- Beperk de datatoegang tot personen die toegang nodig hebben voor het maken van statistieken en beheer en evalueer deze toegangsrechten periodiek.