Ministerie van Justitie en Veiligheid

Minister voor Rechtsbescherming

Directoraat-Generaal Rechtspleging en Rechtshandhaving
Directie Rechtsbestel

Turfmarkt 147
2511 DP Den Haag
Postbus 20301
2500 EH Den Haag
www.rijksoverheid.nl/jenv

nota

Aanbiedingsbrief rapport 'The influence of (technical) developments on the concept of personal data in relation to the GDPR'

Datum
6 december 2022

Ons kenmerk
4362216

1. Aanleiding

Het WODC heeft het op verzoek van JenV uitgevoerde onderzoek 'The influence of (technical) developments on the concept of personal data in relation to the GDPR' opgeleverd.

2. Gevraagde besluit

Instemmen met de verzending van bijgevoegde brief aan de Tweede Kamer, waarin het WODC-rapport 'The influence of (technical) developments on the concept of personal data in relation to the GDPR' aan de Tweede Kamer wordt aangeboden.

3. Kernpunten

Het WODC zal het rapport in ieder geval zes weken na oplevering, dus op 20 december 2022 openbaar maken. Het is te doen gebruikelijk een WODC-rapport voor die tijd aan te bieden aan de Kamer.

4. Toelichting

Aanleiding onderzoek

Op 20 november 2020 schreef de toenmalig minister voor Rechtsbescherming aan de Kamer dat het kabinet voornemens was om een onderzoek uit te laten voeren naar de invloed van de voortgaande technologische ontwikkelingen op het begrip persoonsgegevens.[1] Aanleiding daarvoor was het beleid op het terrein van de regulering van algoritmen waarin de Algemene verordening gegevensbescherming (AVG) een cruciale rol speelt. Besluitvormingsonderzoek en het procesrechtonderzoek indiceerden dat door de technologische ontwikkeling onzekerheid ontstaat over de vraag of de AVG op bepaalde gegevenssets van toepassing is. De onderzoeken kwamen tot verschillende inzichten ten aanzien van de toepasselijkheid van de AVG. Volgens het besluitvormingsonderzoek zullen gegevenssets door de inzet van algoritmische modellen eerder persoonsgegevens bevatten, omdat verbanden in voorheen anonieme gegevenssets kunnen worden gevonden. Het procesrechtonderzoek wijst juist op technische mogelijkheden om persoonsgegevens te anonimiseren, waardoor de AVG niet meer zou gelden. Om tot een helder overzicht van de technische ontwikkelingen en hun juridische impact te komen, besloot het kabinet nader onderzoek laten verrichten. Dat

[1] Kamerstuk 26643, nr. 726, p. 5.

onderzoek is op verzoek van het ministerie van Justitie en Veiligheid door het WODC ter hand genomen en uitgevoerd door het Tilburg Institute for Law, Technology and Society van de Universiteit Tilburg.

In deze nota wordt kort ingegaan op de bevindingen van de onderzoekers, zonder dat daarover een inhoudelijk oordeel wordt gegeven. Daarvoor is nadere ambtelijke bestudering en afstemming nodig. Deze zal de komende maanden plaatsvinden.

Probleemstelling en onderzoeksvraag
De AVG stelt regels aan en bevat normen voor de verwerking van persoonsgegevens. Hoewel de AVG sinds 2018 van kracht is stammen de regels in essentie uit de jaren ’70 van de vorige eeuw. Doorslaggevend voor de toepassing van het gegevensbeschermingskader was toen, en is vandaag de dag nog steeds, of de verwerkte gegevens informatie van een identificeerbaar individu (natuurlijke persoon) betreffen. Hoewel een dergelijke vaststelling in de jaren ‘70 van de vorige eeuw relatief eenvoudig was, is dit in de loop van de tijd steeds complexer geworden, vooral in het licht van technologische ontwikkelingen, de algemene toegankelijkheid van technologieën en de push tot meer open data. Deze ontwikkelingen hebben tot gevolg dat het steeds makkelijker is om persoonsgegevens af te leiden uit datasets die dergelijke gegevens op het eerste gezicht niet lijken te bevatten. Ze hebben ook tot gevolg dat de juridische status van data steeds meer fluïde wordt: doordat data worden gedeeld tussen partijen en de verwerkingen van datasets aanzienlijk verschillen, kan dezelfde dataset het ene moment worden gekwalificeerd als een dataset met persoonsgegevens en het andere moment niet, of als dataset met persoonsgegeven in handen van partij A maar tegelijkertijd zonder persoonsgegevens in handen van partij B.

De vraag of een dataset persoonsgegevens bevat is bepalend bij de beslissing of de gegevensbeschermingsregels van toepassing zijn. De Europese Unie (EU) heeft een ander kader vastgesteld voor de verwerking van niet-persoonsgegevens. In dat geval mogen er in wezen geen beperkingen worden gesteld met betrekking tot het vrije verkeer van niet-persoonsgegevens. De juridische kwalificatie of een dataset al dan niet persoonsgegevens bevat, betekent dus het verschil tussen reguleringskaders die bijna 180 graden van elkaar afwijken.

Bij persoonsgegevens gaat het niet alleen om directe, maar ook om indirecte informatie, dat wil zeggen gegevens waaruit de identiteit van een persoon kan worden afgeleid. Het gaat bovendien niet alleen om identificerende gegevens—gegevens die op dit moment tot een bepaalde persoon kunnen leiden—maar ook om identificeerbare gegevens. Met andere woorden om gegevens die op dit moment niet tot een bepaalde persoon leiden, maar in de toekomst mogelijk wel. Om te bepalen of een dataset identificeerbare gegevens bevat, moet rekening worden gehouden met alle middelen waarvan redelijkerwijs valt te verwachten dat zij worden gebruikt om gegevens aan een persoon te koppelen. Ten slotte is het niet nodig om de identiteit van een persoon te kennen: als gegevens worden gebruikt om een beslissing te nemen over een specifieke persoon wiens identiteit onbekend is, is het gegevensbeschermingsregime al van toepassing.

Er zijn ook belangrijke technologische en maatschappelijke ontwikkelingen van invloed op de vraag of datasets persoonsgegevens bevatten. Big Data, Artificiële Intelligentie, Quantum Computing en andere technieken maken het nog gemakkelijker om persoonsgegevens af te leiden uit geaggregeerde,

geanonimiseerde of versleutelde datasets; de algemene toegankelijkheid van technologieën maakt het nog moeilijker om de toekomstige status van een dataset te bepalen; en het voortdurende streven naar open data en het hergebruik van overheidsinformatie betekent dat de juridische status van data nog meer fluïde zal worden.

Tegen deze achtergrond was de probleemstelling voor dit onderzoek:
Welk effect hebben huidige en toekomstige technische ontwikkelingen op het gebied van anonimisering, pseudonimisering, aggregatie en identificatie van gegevens, op het gegevensbeschermingskader en de bescherming van de verschillende soorten gegevens?

Belangrijkste conclusies
Het zal steeds moeilijker worden om de (juridische) anonimiteit van datasets te waarborgen. Experts die voor dit onderzoek zijn geïnterviewd, betwijfelen nu al of het mogelijk is om aan de wettelijke criteria voor anonimiteit te voldoen. Terwijl het wettelijke regime anonimiteit als een binair vraagstuk beschouwt, zien de meeste technische experts het als een schaal. De meeste technologieën en tegen-technologieën zijn verwikkeld in een kat-en-muisspel. Dit wordt ook verondersteld het geval te zijn voor de toekomst van onder meer anonimiserings- en de-anonimiseringstechnieken, aggregatie- en inferentietechnieken en voor encryptie en decryptie. De meest fundamentele verschuiving is de algemene toegankelijkheid van dergelijke technologieën. Dit betekent dat, vooral wanneer gegevens online beschikbaar worden gesteld, het steeds waarschijnlijker wordt dat er wereldwijd enkele partijen zullen zijn die geavanceerde technologieën zullen gebruiken om gegevens te ontsleutelen, opnieuw te identificeren of te de-anonimiseren en de nodige tijd, energie en moeite te investeren om dit doen.

De algemene beschikbaarheid van open data en de algemene toegankelijkheid van datatechnologieën leiden volgens de onderzoekers tot een aantal fundamentele vragen met betrekking tot aspecten van de AVG.

Ten eerste is de aard van de data in Big Data-processen niet stabiel, maar volatiel. Een dataset met gewone persoonsgegevens kan worden gekoppeld aan en verrijkt met een andere dataset om gevoelige gegevens af te leiden; de gegevens kunnen vervolgens worden samengevoegd of ontdaan van identificatiegegevens en niet-persoonlijk worden; vervolgens kunnen de gegevens worden gedeanonimiseerd of geïntegreerd in een andere dataset om opnieuw persoonsgegevens te creëren. Dit alles kan in een fractie van een seconde gebeuren. De vraag is daarom of het zin heeft om met afgebakende categorieën te werken als dezelfde dataset letterlijk van de ene seconde op de andere in een andere categorie kan vallen en de volgende seconde in weer een andere.

Ten tweede wordt het als gevolg van het voorgaande steeds moeilijker om de status van gegevens precies te bepalen. Om de huidige status van een dataset te bepalen, moet rekening worden gehouden met de verwachte toekomstige status van de gegevens. Gezien de toenemende toegankelijkheid van technologieën wordt het steeds waarschijnlijker dat wanneer een database wordt gedeeld of anderszins beschikbaar wordt gesteld, er een partij is die hieruit persoonsgegevens destilleert, maar ook dat er andere partijen zijn die dat niet willen. De juridische categorie waartoe de gegevens behoren, is dus niet langer een kwaliteit van de gegevens zelf, maar een product van de inspanningen en investeringen van een verwerkingsverantwoordelijke. Bijgevolg is het de vraag of

anonimiseren of pseudonimiseren kan worden bereikt in een context waarin het bepalen van de status van gegevens nauwelijks haalbaar is.

Ten derde zijn moderne gegevensverwerkingen in toenemende mate gebaseerd op geaggregeerde gegevens, die ook zeer grote individuele en sociale gevolgen kunnen hebben. Het profileren van doelgroepen in plaats van individuen wordt een gangbare verwerkingshandeling in de informatiemaatschappij. De gevolgen van deze activiteiten kunnen negatief zijn voor de individuen in een groep, zonder dat de schade direct te relateren is aan individuen. Het idee dat hoe gevoeliger de gegevens zijn en hoe directer ze aan een persoon kunnen worden gekoppeld, des te strikter de verwerking ervan moet worden gereguleerd, kan daarom in twijfel worden getrokken.

De inzichten uit dit onderzoek dienden te worden gerelateerd aan de vraag in hoeverre er een situatie van onder- of overregulering kan ontstaan wanneer geanonimiseerde of herleide gegevens niet, respectievelijk wel onder de reikwijdte van de AVG vallen. De onderzoekers beargumenteren dat het beantwoorden van deze vragen afhangt van wat wordt beschouwd als de reguleringsdoelstelling van het gegevensbeschermingsregime:

- Moet het gegevensbeschermingskader worden beschouwd vanuit een beschermend perspectief of vanuit het perspectief van het faciliteren van gegevensverwerking binnen een vastgesteld kader, of als een combinatie tussen beide?
- Moet het beschermingsdoel worden opgevat als het voornamelijk bieden van bescherming aan individuele belangen of (ook) aan groeps- en maatschappelijke belangen?
- Moet het gegevensbeschermingsregime worden opgevat als het stellen van beperkingen voor gegevensverwerking of als het bieden van een kader voor het gebruik en het delen van gegevens?
- Is de beschermende grondgedachte het best gediend door beperkingen, of kan er soms meer gegevensverwerking nodig zijn om de belangen van individuen en/of de samenleving zo goed mogelijk te dienen?
- Is de grondgedachte van het faciliteren van gegevensgebruik het best gediend met een open en contextueel kader of met het stellen van strikte en duidelijke regels waarbinnen gegevensverwerking als legitiem wordt beschouwd?

Dit onderzoek had niet ten doel om op deze vragen een definitief antwoord te geven; wel is duidelijk dat afhankelijk van de antwoorden op deze vragen verschillende lacunes in de regelgeving en gevaren voor over- en/of onderregulering zullen worden gevonden.

Overkoepelende analyse
Het is duidelijk dat de technologische ontwikkelingen en de algemene beschikbaarheid van data nu en in de toekomst tot gevolg hebben dat anonimiseren steeds moeilijker wordt. De status van data wordt steeds volatieler en wordt steeds minder een kenmerk van data en datasets zelf en meer en meer een effect van de inspanningen van de verwerkingsverantwoordelijke. De juridische categorieën zullen steeds meer fluïde en minder stabiel worden en één database kan juridisch verschillend worden beoordeeld ten aanzien van verschillende partijen die daar toegang toe hebben.

Gezien deze trends en gezien de begrippen ‘identificeerbaarheid’ en ‘alle middelen waarvan redelijkerwijs valt te verwachten dat zij worden gebruik’, zullen steeds meer gegevens, zo niet alle, onder het gegevensbeschermingskader vallen.

Directoraat-Generaal Rechtspleging en Rechtshandhaving
Directie Rechtsbestel

Datum
6 december 2022

Ons kenmerk
4362216

Het onderzoek laat voor wat betreft de technologische ontwikkeling een eenduidig scenario zien: anonimiseren wordt steeds moeilijker, juridische categorisering zal steeds moeilijker worden en de status van gegevens zal steeds meer een effect zijn van de inspanningen van de verwerkingsverantwoordelijke. De onderzoekers beschrijven vijf globale strategieën waarmee het wettelijk regime zou kunnen reageren op de toegenomen beschikbaarheid van open data en de algemene toegankelijkheid van technologie:

- Het huidige gegevensbeschermingskader intact laten,
- Focussen op duidelijkere gegevenscategorieën,
- Meer nadruk leggen op contextualiteit,
- Gebruikmaken van verschillende gegevenscategorieën en daaraan gekoppelde reguleringsregimes,
- Focussen op een volledig contextueel gegevensbeschermingskader.

Taal
Vanwege de internationale relevantie van de thematiek is in afstemming met het WODC en na advies van de begeleidingscommissie besloten het eindrapport in het Engels te schrijven. De conclusies van het onderzoek zijn immers op een Europees niveau van belang en eventuele discussies naar aanleiding van de conclusies van dit onderzoek moeten op Europees niveau worden gevoerd. Voor de Nederlandse context is een uitgebreide en zelfstandig leesbare samenvatting van het eindrapport in het Nederlands geschreven.

Vervolg
Op 20 december 2022 plaatst het WODC het rapport onder begeleiding van een nieuwsbericht op zijn website en wordt het rapport openbaar. U wordt derhalve geadviseerd het rapport voordien aan de Kamer aan te bieden. Aan het slot van de aanbiedingsbrief kondigt u aan in het eerste kwartaal van 2023 te komen met een kabinetsreactie op het onderzoek. Tot dat moment wordt u geadviseerd desgevraagd te melden dat het rapport en de aanbevelingen worden bestudeerd.

5. Informatie die zich niet leent voor openbaarmaking
De persoonsgegevens van de ambtenaren zijn niet openbaar ter bescherming van de persoonlijke levenssfeer.