Bijlage 2 bij Privacyreglement NIVEL Zorgregistraties eerste lijn

Technische beschrijving pseudonimisatie gegevensverzameling NIVEL Zorgregistraties eerste lijn

Pseudonimisatie

Onder 'pseudonimisatie' verstaan wij het omzetten van een persoonsgegeven naar een niet-herleidbare code. De omzettingen zijn, in de door ZorgTTP gehanteerde variant, onomkeerbaar. Het is daarbij onmogelijk een pseudoniem terug te vertalen naar het oorspronkelijke persoonsgegeven.

De omzetting verloopt in twee stappen: de partij die in het bezit is van de te verzenden (persoons)gegevens (de bron) maakt gebruik van pseudonimisatiesoftware waarmee een persoonsgegeven wordt omgezet naar een zogenaamd pre-pseudoniem. Volgens wordt als tweede stap in het proces het pre-pseudoniem door de TTP, met behulp van software, omgezet naar een definitief pseudoniem. Dit pseudoniem, en de bijbehorende overige data, worden beschikbaar gesteld aan de ontvangende partij (het doel).

Alleen de TTP weet op welke wijze het definitieve pseudoniem is aangemaakt. Daarmee wordt een situatie bereikt waarbij het voor zowel de bron als het doel (de ontvangende partij) onmogelijk is om het oorspronkelijke persoonsgegeven met het aangemaakte pseudoniem in verband te brengen.

Het College Bescherming Persoonsgegevens (CBP) stelt de volgende eisen aan een TTP die pseudonimisatie dienstverlening aanbiedt[1]:

- Er wordt vakkundig gebruik gemaakt van pseudonimisering, waarbij de eerste pseudonimisatie plaatsvindt bij de aanbieder;
- Er zijn technische en organisatorische maatregelen genomen om herhaalbaarheid van de versleuteling (‘replay attack’) te voorkomen;
- De verwerkte gegevens zijn niet indirect identificerend;
- In een onafhankelijk deskundig oordeel (audit) wordt vooraf en daarna periodiek vastgesteld dat aan voorwaarden 1), 2) en 3) is voldaan;
- De werkwijze wordt beschreven en gepubliceerd.

Voor het pseudonimiseren van privacygevoelige gegevens heeft ZorgTTP een pseudonimisatieplatform ontwikkeld. Dit omvat naast de ontwikkelde software ook technische en organisatorische voorzieningen om zorgvuldig alle werkzaamheden te kunnen uitvoeren.

Het pseudonimisatieplatform wordt jaarlijks onderworpen aan zowel interne als externe audits. De externe audit wordt uitgevoerd door Pricewaterhouse Coopers (PwC). Deze audit wordt, conform door het CBP gestelde eisen, jaarlijks herhaald om de vorderingen te beoordelen. De door ZorgTTP gehanteerde werkwijze is daarbij steeds positief beoordeeld.

De kerntaak van ZorgTTP is het depersonaliseren van bestanden om daarmee het uitwisselen van informatie op individueel niveau, conform de wettelijke vereisten, mogelijk te maken. De verzendende (de bron) en de ontvangende partij (het doel) maken gezamenlijk afspraken over welke informatie wordt uitgewisseld en welke gegevens daarbij dienen te worden geanonimiseerd. ZorgTTP zal een adviserende rol spelen bij deze afweging als onderdeel van de werkzaamheden die horen bij het inrichten van een pseudonimisatieketen. Het is echter aan het CBP, als toezichthouder op de Wet Bescherming Persoonsgegevens (WBP), voorbehouden om de gemaakte afwegingen te toetsen en een oordeel te vellen over de mate waarin de uitwisseling is toegestaan binnen de kaders die de WBP stelt.

[1] Brief CPB aan Ministerie van VWS, 6 maart 2007; kopie bij ZorgTTP aanwezig.

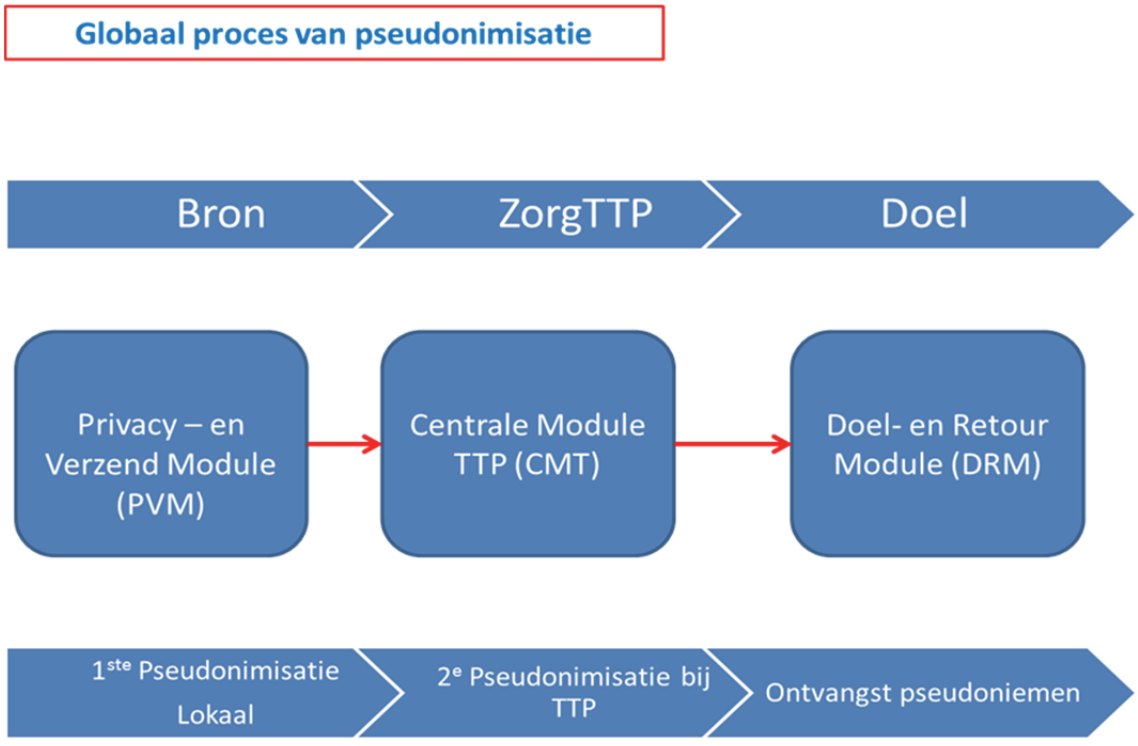
Bij ZorgTTP wordt geen data, zoals patiëntbestanden, opgeslagen. Enkel het algoritme waarmee de persoonsgegevens worden gepseudonimiseerd wordt bewaard.

Beknopte beschrijving van het pseudonimisatieproces

Het pseudonimisatieproces bestaat in het kort uit de volgende stappen:

1. Uitgangspunt is dat de verzendende partij (de zorgverlener) een bestand genereert dat voldoet aan vooraf vastgestelde specificaties;
2. Het bestand wordt verwerkt met de door ZorgTTP aan de databron beschikbaar gestelde software;
3. Na verwerking volgt beveiligd transport naar ZorgTTP voor het aanmaken van de definitieve pseudoniemen;
4. ZorgTTP voert met behulp van eigen pseudonimisatiesoftware centraal een tweede bewerking uit waarbij een voor de zender en ontvanger geheime ‘sleutel’ wordt gebruikt;
5. Het gepseudonimiseerde bestand wordt vrijgegeven en kan vervolgens worden opgehaald door de ontvangende partij met een daartoe beschikbaar gestelde ontvangstmodule.

In onderstaande figuur wordt het pseudonimisatieproces schematisch weergegeven.



Privacy- en Verzend Module (PVM)

Deze module wordt gebruikt door de bron en kent een aantal functies. Allereerst wordt een aantal controles uitgevoerd op de aangeboden gegevens. Daarna worden de identificerende persoonsgegevens omgezet in zogenaamde pre-pseudoniemen.
Pre-pseudoniemen zijn persoonsgegevens waarop een eerste bewerking heeft plaatsgevonden. Vervolgens wordt een scheiding aangebracht tussen de pseudoniemen (sleuteldeel) en de bijbehorende data (datadeel). Beide delen worden vervolgens geëncrypteerd. Het sleuteldeel kan enkel worden gedecrypteerd door ZorgTTP, het datadeel enkel door de uiteindelijke ontvanger.
Het uitvoeren van een eerste bewerking op de te pseudonimiseren gegevens bij de bron is een vereiste binnen het kader dat het CBP heeft opgesteld.
Voordat van een onomkeerbaar pseudoniem gesproken kan worden, dient de TTP een definitieve omzetting te doen op de voorbewerkte gegevens. De gegevens worden op beveiligde wijze naar de TTP verstuurd. Daarbij zijn de gegevens zodanig beveiligd dat deze slechts voor de TTP toegankelijk zijn voor verdere bewerking. De hiertoe benodigde op Java gebaseerde software wordt via het internet beschikbaar gesteld en maakt gebruik van door de TTP uitgegeven digitale certificaten. De digitale certificaten worden gebruikt voor ondertekening van de te verzenden berichten, het opbouwen van een beveiligde (HTTPS-) verbinding en encryptie van de te verzenden data.

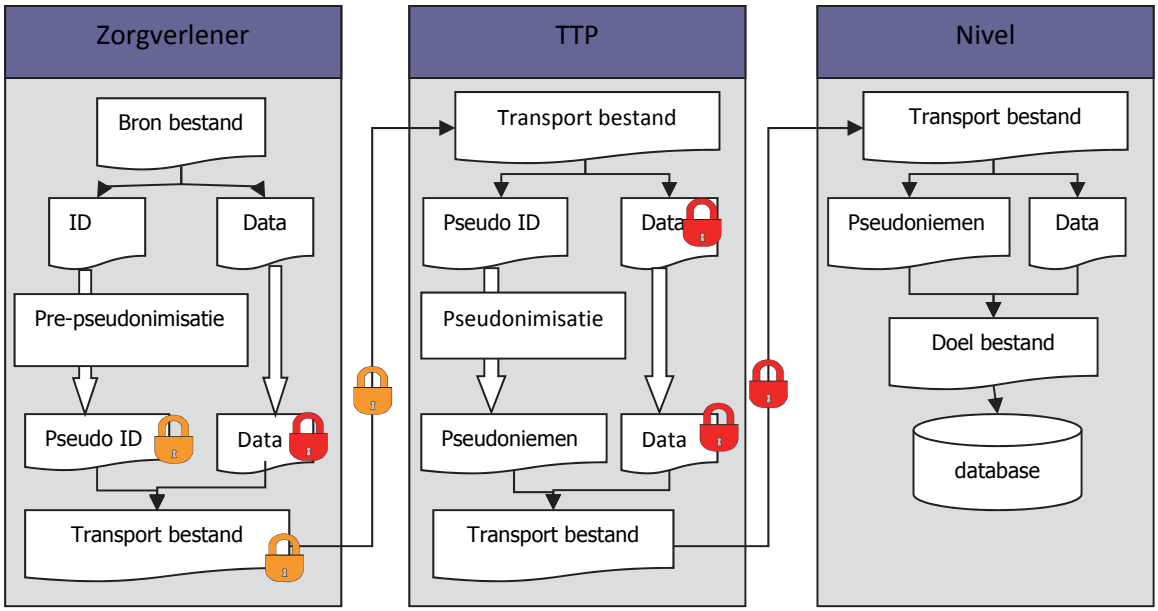
Centrale Module TTP (CMT)

De centrale applicatie ontvangt een versleuteld bestand. Dit bestand bestaat uit twee onderdelen: een datadeel en een sleuteldeel. Het sleuteldeel bevat de pre-pseudoniemen, deze worden door de centrale applicatie omgezet tot de definitieve pseudoniemen.

De centrale applicatie heeft geen toegang tot het datadeel, deze is beveiligd en enkel door de ontvangstapplicatie te decrypteren. Voor de transportbeveiliging wordt ook weer gebruik gemaakt van een Public Key Infrastructure (PKI).

Doel- en Retour Module (DRM)

De ontvangstmodule wordt gebruikt door de ontvangende partij. De module ontvangt van de centrale applicatie de berichten. De berichten hebben een multipart-xml-indeling. Het is feitelijk een container met daarin bestanden. De module ontsleutelt allereerst het sleuteldeel, vervolgens het datadeel en voegt deze daarna weer samen. Afhankelijk van de grootte van het bestand kost dit proces enkele seconden tot een minuut.



Herleidbaarheid gepseudonimiseerde gegevens

Volgens de CBP uitgangspunten wordt gepseudonimiseerde data niet langer beschouwd als tot persoonsgegevens terug te herleiden data. De opgeslagen gegevens blijven echter een zekere mate van gevoeligheid behouden, zeker in het geval van medische gegevens. Daarnaast kan door het koppelen van gepseudonimiseerde dataverzamelingen of door het toevoegen van aanvullende variabelen alsnog op directe of indirecte wijze sprake zijn van tot persoonsgegevens herleidbare data.

In dat geval zal de registratie alsnog dienen te voldoen aan de eisen die de WBP stelt aan het verwerken van persoonsgegevens.

Om de kans op indirecte herleidbaarheid te minimaliseren adviseert ZorgTTP om:

- gegevens waar mogelijk op geaggregeerd niveau te verstrekken;
- per gepseudonimiseerde dataverzameling met een andere geheime sleutelwaarde te werken. Daarmee wordt directe koppeling op grond van de pseudoniemen onmogelijk;
- gepseudonimiseerde data op het laagste aggregatieniveau alleen op basis van een overeenkomst te verstrekken;
- gepseudonimiseerde data op het laagste aggregatieniveau uit andere gepseudonimiseerde dataverzamelingen alleen toe te voegen na analyse van het risico op directe op indirecte herleidbaarheid.

Historische data

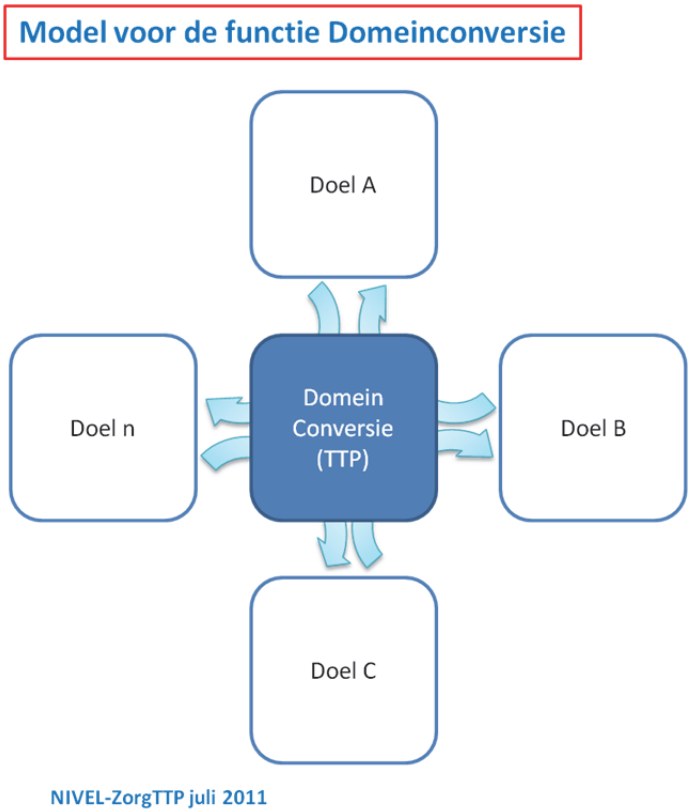
Het NIVEL verzamelt al geruime tijd medische informatie. Zorgregistratie huisartsen (LINH) is bijvoorbeeld een netwerk van 84 geautomatiseerde huisartspraktijken met meer dan 335.000 ingeschreven patiënten (juli 2010). De LINH-huisartsen verzamelen op continue basis 'productiegegevens' over aandoeningen (ICPC-gecodeerde diagnose), aantallen contacten/verrichtingen, geneesmiddelvoorschriften en verwijzingen. LINH is vanaf 1991 operationeel.
Vanaf medio 2012 wordt gefaseerd het pseudonimisatieplatform in gebruik genomen. Daarbij wordt de historische dataverzameling ook gepseudonimiseerd. Daardoor zijn de nieuwe gepseudonimiseerde aanleveringen te koppelen aan de historische dataverzameling.

Domeinconversie

ZorgTTP verleent pseudonimisatiediensten aan diverse partijen. Deze en andere partijen die al dan niet over gepseudonimiseerde dataverzamelingen beschikken, hebben behoefte aan de mogelijkheid om – op gecontroleerde wijze – bestanden aan elkaar te kunnen koppelen. Daarom is een functie voor zogenaamde ‘domeinconversie’ ontwikkeld.

Domeinconversie maakt het mogelijk om een pseudoniem van een BSN van het ene domein (lees een gepseudonimiseerde dataverzameling) te converteren naar een pseudoniem, op grond van hetzelfde BSN, zoals bekend binnen een ander domein.
Het reguliere pseudonimisatieproces verloopt in twee stappen. De eerste versleuteling vindt plaats bij de bron, de tweede versleuteling bij ZorgTTP. Onderdeel van de tweede versleuteling is een domeinspecifieke encryptie. Dit betekent dat iedere gepseudonimiseerde dataverzameling van een specifieke serie pseudoniemen wordt voorzien. De kracht van deze domeinspecifieke encryptie is dat kan worden voorkomen dat gepseudonimiseerde dataverzamelingen eenvoudigweg op basis van pseudoniemen aan elkaar gekoppeld kunnen worden. Daarmee zou het risico op indirecte herleidbaarheid naar de oorspronkelijke persoonsgegevens onaanvaardbaar hoog worden.

Voor elke opdrachtgever kan daardoor voorzien worden in één of meer domeinen.
In de praktijk betekent dit dat gelijke input – bijvoorbeeld een bepaald BSN – in de verschillende domeinen verschillende pseudoniemen zal opleveren. Slechts met behulp van domeinconversie kunnen verschillende domeinen aan elkaar gekoppeld worden.



Communicatiemodel

Aangezien het in specifieke gevallen mogelijk moet zijn om additionele informatie te verzamelen, wordt er gebruik gemaakt van een communicatiemechanisme. De reguliere pseudonimisatie wordt zoals eerder beschreven uitgevoerd. Voor het verzoek om additionele informatie zijn twee zaken noodzakelijk: een identificatie van de behandelaar en een lokaal patiëntnummer.

Als de onderzoeker tot de conclusie komt dat additionele informatie van grote meerwaarde zou zijn, dan kan gebruik worden gemaakt van de communicatiedatabase. Het is dan wel noodzakelijk dat er een gedeelde variabele is tussen beide domeinen. We maken daarvoor gebruik van een communicatiepseudoniem. Dit pseudoniem is zowel beschikbaar in het domein Onderzoek als in het domein Communicatie. Omdat het twee domeinen betreft, zal het pseudoniem – ondanks identieke input – een andere waarde opleveren. ZorgTTP kan het communicatiepseudoniem – onder strikte voorwaarden - converteren van het domein Onderzoek naar het domein Communicatie.

Het proces verloopt dan als volgt:

1. NIVEL bepaalt in welke gevallen er sprake is van grote meerwaarde van additionele informatie. Enkel in die gevallen wordt er gebruik gemaakt van het communicatiemechanisme;
2. De aanvraag betreft een zogenaamde ‘domeinconversie’. Deze moet technisch door twee medewerkers van ZorgTTP worden goedgekeurd;
3. Vanuit NIVEL wordt de zorgverlener gevraagd om de patiënt te benaderen met het verzoek om additionele informatie;
4. De patiënt beslist zelf of hij of zij de informatie verstrekt aan NIVEL;
5. Bij het verstrekken van de informatie wordt gewerkt met een onderzoeksnummer, er zijn geen persoonsgegevens noodzakelijk.

Schematische weergave communicatiemodel