

Aanleverspecificaties bestand herpseudonimisatie CBS

Inleiding

Dit document betreft de aanleverspecificaties voor het aanbieden van pseudoniemen aan de PVM CBS voor herpseudonimisatie (oftewel omzetting naar CBS-pseudoniemen).

De Privacy Verzend Module (PVM) leest de te pseudonimiseren gegevens uit een bestand in CSV formaat met de volgende kenmerken:

- Het scheidingsteken is de puntkomma (;)
- Het quote-teken is het dubbele aanhalingsteken ("")
- De tekenset van het bestand is ISO-8859-1
- De bestandsnaam dient te eindigen op de extensie .csv

Bestandsnaam

De bestandsnaam is vrij te kiezen, zodanig dat betekenis aan de zending omtrent inhoud, afzender, datum en gebruik zijn toe te kennen. ZorgTTP raadt aan een conventie te hanteren (zonder dat de PVM deze afdwingt of zich hiertoe beperkt): `cbs_test` of `productie_afzender_datum_volgnummer.csv`, bijvoorbeeld:

Naam	Betekenis
<code>cbs_prod_\$afzender\$_20180801_1.csv</code>	Productie aanlevering van \$afzender\$ op 1 augustus 2018, 1 ^e bericht
<code>cbs_test_\$afzender\$_20180201_2.csv</code>	Test bericht van \$afzender\$ op 1 februari 2018, 2 ^e poging

Opbouw bestand

Voor de verwerking van bestanden maakt de PVM gebruik van kolomlabels. De volgende kolomlabels worden als input gebruikt voor herpseudonimisatie: **PSEUDONIEM BSN**, **PSEUDONIEM CBS** en **PSEUDONIEM NIVEL**. Alle andere variabelen worden ongewijzigd aangeboden aan het CBS.

Kolom	Criteria
PSEUDONIEM BSN, CBS of NIVEL	een ZorgTTP pseudoniem, bestaande uit een prefix van 4 karakters en een waarde van 28 karakters. Alle pseudoniemen in deze kolom moeten afkomstig zijn uit hetzelfde domein.

De informatie uit het inputbestand in andere dan de genoemde, voor pseudonimisatie relevante kolom wordt versleuteld door de keten getransporteerd en *onveranderd* bij het CBS afgeleverd.

Voorbeeld input herpseudonimisatie CBS:

```
VOLGNR;PSEUDONIEM BSN;PSEUDONIEM CBS;PC4;DATONTVANGST;LIS_BIOPT
1; TP2B0001g9davihecoh429lot182t6gb;TP2C0027kd9sfi1t89mp6vjdn79gj71c;3995;12-10-2018;LIS
```

Waarbij de PVM een bestand alleen verwerkt wanneer de eerste regel kolomnamen bevat en de kolom 'PSEUDONIEM BSN', 'PSEUDONIEM CBS' of 'PSEUDONIEM NIVEL' daarin voorkomt.