

To: [REDACTED] 5.1.2e [REDACTED] 5.1.2e @rivm.nl]; [REDACTED] 5.1.2e [REDACTED] 5.1.2e @rivm.nl]
From: [REDACTED] 5.1.2e
Sent: Fri 3/5/2021 10:27:09 AM
Subject: RE: de weg terug, back to where we came from
Received: Fri 3/5/2021 10:27:10 AM

Dank, [REDACTED] 5.1.2e Gelukkig!

From: [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>
Sent: vrijdag 5 maart 2021 11:26
To: [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>; [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>
Subject: FW: de weg terug, back to where we came from

Hoi,

Zie mailwisseling hieronder. Ik had het gisteren met [REDACTED] 5.1.2e over koppelen van CIMS data.

Er lijken dus wel mogelijkheden, gelukkig.

Gr., [REDACTED] 5.1.2e

From: [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>
Sent: vrijdag 5 maart 2021 10:39
To: [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>; [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>; [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>; [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>; [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>
Subject: RE: de weg terug, back to where we came from

Hoi [REDACTED] 5.1.2e

Over 'toegestaan' moet de commissie oordelen die hiervoor ingericht worden. Het is wel mogelijk. De depseudonimiseringfunctie kan daartoe bij een geselecteerde dataset:

- de pseudoniemen vervangen door de echte identificerende kenmerken (BSN en/of clientnummer) en
- de dataset aanvullen met benodigde (en toegestane) persoonsgegevens uit CIMS.

Uitgangspunt is dat deze dataverwerking in de depseudonimiseringfunctie plaatsvindt onder een aparte rol en autorisatie, na goedkeuring door de voornoemde commissie.

Met vriendelijke groet,

[REDACTED] 5.1.2e
[REDACTED] 5.1.2e

Rijksinstituut voor Volksgezondheid en Milieu
Ministerie voor Volksgezondheid, Welzijn en Sport
A. van Leeuwenhoeklaan 9 | 3721 MA | Bilthoven
Postbus 1 | 3720 BA Bilthoven

M: [REDACTED] 5.1.2e
E: [REDACTED] 5.1.2e @rivm.nl

Van: [REDACTED] 5.1.2e <[REDACTED] 5.1.2e @rivm.nl>
Verzonden: vrijdag 5 maart 2021 09:41

Aan: 5.1.2e <5.1.2e@rivm.nl>; 5.1.2e <5.1.2e@rivm.nl>; 5.1.2e
<5.1.2e@rivm.nl>; 5.1.2e <5.1.2e@rivm.nl>

Onderwerp: RE: de weg terug, back to where we came from

Hoi,

vanuit de inhoud dan een aantal voorbeelden, waarbij er (vgs mij) noodzaak is tot depseudonimiseren.

Voorbeelden moeten dus noodzaak hebben om identificerende gegevens te hebben:

1. Recall; voorbeeld: er is gevaccineerd met vaccin&batch, waarbij opvallend veel van een specifieke ernstige bijwerking worden gemeld (vb overlijden). Dan wil je de groep mensen, die hiermee gevaccineerd is, hierover berichten.
2. Een groep onderzoekers in Noord Nederland wil onderzoek gaan uitvoeren met mensen, die wel en niet zijn gevaccineerd tegen covid. Zij willen die mensen uitnodigen en toestemming vragen voor deelname aan de studie. Dit is een data-aanvraag van een externe partij, die beoordeeld zal worden door een nog op te zetten commissie.
3. Het RIVM wil meer inzicht krijgen in lange termijn risicofactoren van covid vaccinatie. Zij willen de gegevens van alle mensen in CIMS koppelen aan andere registraties, die bij CBS beschikbaar zijn. CBS koppelt dan de verschillende registraties door ze allemaal op dezelfde manier te pseudonimiseren en een unieke koppelvariabele te geven. Dit is echt verdiepend onderzoek en zal waarschijnlijk ook weer moeten worden voorgelegd aan de commissie, die CIMS aanvragen gaat beoordelen.
4. Het RIVM wil de effectiviteit van covid19 vaccinatie tegen ziekenhuisopname en sterfte berekenen. Hiervoor is het nodig om de CIMS gegevens te koppelen aan ziekenhuisopname data en sterfte data. Deze koppeling zal worden uitgevoerd door Trusted Third Party. Dit hoort echt bij de monitoring van de vaccinatie campagne en is dus taak, die VWS aan RIVM-EPI heeft gegeven.

Uitdaging is dat voor aantal voorbeelden er koppeling met andere registraties nodig is en ik heb gehoord dat koppeling niet is toegestaan? Ik hoop dat ik het verkeerd heb gehoord/begrepen, maar ik heb daarom 5.1.2e even in de CC gezet.

Vraag vanuit mijzelf is toch nog steeds: als je nu identificerende gegevens nodig hebt voor iets, waarom dan niet rechtstreeks een query op de CIMS database zelf uitvoeren? Dat zal ongetwijfeld te simpel gedacht zijn en te maken hebben met ICT technische zaken*

Gr., 5.1.2e

From: 5.1.2e <5.1.2e@rivm.nl>

Sent: woensdag 3 maart 2021 08:45

To: 5.1.2e <5.1.2e@rivm.nl>

Cc: 5.1.2e <5.1.2e@rivm.nl>

Subject: FW: de weg terug, back to where we came from

Dag 5.1.2e

Hierbij eerste aanzet van pseudonimiseren vanuit de techniek gezien (A) ..

De vraag is hoe het pseudonimiseren in de organisatie processen gaan (B)...

A en B moeten op elkaar aansluiten *

Groet, 5.1.2e

From: 5.1.2e
Sent: dinsdag 2 maart 2021 16:29
To: 5.1.2e <5.1.2e@rivm.nl>
Subject: de weg terug, back to where we came from

DEPSEUDONIMISEREN

De pseudonimasatie en depseudonimisatie geschiedt met functionaliteit in de BICIMS-database, in schema CIMS_KEYS. De techniek bestaat uit de functie GRIP_HASH en een tabel GRIP_HASH_TAB met versleutel-keys. De versleutel-keys zijn voor iedereen onzichtbaar.

Via de views STG_xx_V in schema CIMS_STAGE wordt middels de functie cims_keys.grip_hash(",","atribuut) bepaalde attributen gehashed of gepseudonimiseerd doorgegeven aan de STG-tabellen. Deze STG-tabellen bevatten anonieme data welke vervolgens doorgegeven worden aan de CIMS_DWH. De virtualisatie-laag TDV haalt vervolgens deze data op voor rapportage doeleinden.

Echter is de weg terug vanuit TDV ook wenselijk: Stel dat er wat mis is met een vaccinatie-batch, dan wil men met een bepaalde set bsn-nummers terug kunnen naar de bron CIMS.

Het idee is dan om met TDV een set van anonieme data samen te stellen, in een bepaald csv-formaat om deze vervolgens op een speciale IN-directory te plaatsen waar GRIP/BICIMS bij kan. GRIP herkent het bestand en start daarbij de bijbehorende ETL-flow op die het bewuste csv-formaat kent en het bsn_nummer en/of clientnummer depseudonimiseert. De output van de ETL-flow, de gedepseudonimiseerde data, kan vervolgens in een csv-bestand weggeschreven worden in een speciale UIT-directory, om vervolgens opgepakt te worden door een speciale functionaris.

Middels de file-name pre-fix, bv 'AN_RECALL_02_03_2021.csv' wordt een speciale bijbehorende ETL-FLOW gestart met zijn eigen csv-formaat. Een andere file-name, bv 'AN_MAIL_02_03_2021.csv' heeft zijn specifieke FLOW-ETL en csv-bestand.

De aangeboden anonieme data wordt in schema CIMS_ADATA geladen en geprocessed, gescheiden van andere data. In CIMS_ADATA kan de geanonimiseerde data ook extra verrijkt worden met data van CIMS_DWH, zodat CIMS mogelijk NIET meer benaderd hoeft te worden omdat BICIMS ook alle data wel bevat.

Ook is mogelijk om in deze flows een mail-notificatie te laten versturen naar specifieke adressen, met of zonder de csv als bijlage. De gedepseudonimiseerde data kan overigens ook beschikbaar gesteld worden aan TDV maar de vraag is of dat wenselijk is ?

