

EASHW TIPS – ANONIMITEIT EN PSEUDONIMITEIT

Iedere onderzoeker garandeert vertrouwelijkheid en streeft naar anonimiteit.

4. HERKENBAARHEID: IS IDENTIFICATIE OF RE-IDENTIFICATIE VAN DEELNEMERS MOGELIJK IN DE ANALYSE- DATA EN/OF RESULTATEN?

Er is een verschil tussen (niet-)anoniem onderzoek en (niet-)anonieme data. Er is ook een verschil tussen niet-anonieme, pseudonieme en anonieme data. Deze verschillen zijn belangrijk.

Voor het verschil tussen anonimiteit en pseudonimiteit vertrekken we vanuit de mogelijke herkenbaarheid van deelnemers. We gebruiken ook de termen *identificatie* (deelnemers zijn herkenbaar) en *re-identificatie* (deelnemers kunnen opnieuw herkenbaar gemaakt worden).

ANONIEM & NIET- ANONIEM ONDERZOEK

Onderzoek is anoniem als “niemand door ook maar één iemand herkenbaar is met middelen die redelijkerwijs ingezet kunnen worden”

- Ook als maar 1 iemand een deelnemer kan herkennen is de studie niet anoniem. Dat wil zeggen dat zodra ook maar 1 onderzoeker een deelnemer kan herkennen, er geen sprake kan zijn van anoniem onderzoek. Ook hier is de grens niet helemaal zwart/wit: straatonderzoek, zoals het afnemen van een snelle survey, kunnen we wel als anoniem beschouwen als de onderzoeker geen deelnemers (her)kent bij deze afnamen.
- *Met middelen die redelijkerwijs ingezet kunnen worden:* al wie met bijzondere apparatuur of handelingen wel individuen kan traceren in databestanden worden hier niet meegerekend.

Onderzoek is NIET anoniem als (lijst is niet exhaustief):

- Onderzoekers in contact komen met deelnemers.
- Opnamen gemaakt worden waarbij het gelaat en/of de stem aanwezig zijn (gebruik van foto's, audio- of video-opnamen).
- Opnamen verzameld worden waarbij het gelaat en/of de stem aanwezig zijn (bv. photovoice studies)
- Contactgegevens opgevraagd worden.
- (Andere) identiteitsgegevens opgevraagd worden.
- Data of combinatie van data kunnen leiden tot het herkennen van een deelnemer.

NIET-ANONIEME, ANONIEME EN PSEUDONIEME DATA

Data zijn **niet-anoniem** als men **personen direct of indirect kan herkennen in de dataset**.

Dit is o.a. het geval voor volgende types gegevens, data:

- *Directe herkenbaarheid:* Namen, contactgegevens, audio- of beeldmateriaal en dergelijke.
- *Indirecte herkenbaarheid:* op basis van de combinatie van enkele gegevens. Bijvoorbeeld, de naam van een bedrijf en de functie van een werknemer, of de leeftijd van een *outlier* in een populatie, etc. We vragen jou in te schatten of deelnemer herkenbaar kunnen zijn op basis van de data die je verzamelt.
- Bij twijfel beschouw je de data als niet-anoniem.

Data zijn **anoniem** als men **niemand rechtstreeks of onrechtstreeks kan herkennen, noch kan her-identificeren**.

Dit is het geval indien:

- Je nooit gegevens verzameld hebt waarmee men direct of indirect iemand kan herkennen, OF
- Je *anonimiseert je data*: Je hebt wel gegevens verzameld waarmee men deelnemers kan herkennen, maar je wist deze gegevens definitief. (Bijvoorbeeld audio-opnames vernietigen na transcriptie waarbij deelnemers niet meer herkenbaar zijn in de transcripten).

Data zijn **pseudoniem** als men **personen kan her-identificeren**.

Dit is het geval indien *je data pseudonimiseert* door te werken met:

1. Een databestand waarin deelnemers herkenbaar zijn, en
2. Een databestand waarin deelnemers onherkenbaar gemaakt zijn met een code, en
3. Een sleutelbestand dat de code uit bestand 2 kan koppelen aan de gegevens uit bestand 1.

EASHW ADVISEERT:

- Voorkeur te geven aan anonieme dataverzameling als dit mogelijk is,
- Voorkeur te geven aan het anonimiseren van data boven het pseudonimiseren,
- De termen correct en consequent te gebruiken doorheen je aanvraag,
- Je deelnemers correct te informeren:
 - Is je dataverzameling anoniem of niet?
 - Indien niet: zal je data anonimiseren, pseudonimiseren of niet-anoniem behouden? Leg uit waarom je voor welke optie kiest en vanaf welk moment data al dan niet anoniem/pseudoniem worden.

VISUEEL – DATA PSEUDONIMISEREN VS. ANONIMISEREN

Dit is een visuele weergave op basis van een fictief voorbeeld:

DATASET

Volledige dataset

Naam
E-mail
Werkgever + functie
Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slaap
Gemiddeld aantal publicaties per jaar

Data waarmee deelnemers herkenbaar zijn:

Naam
E-mail
Werkgever + functie

Data waarmee deelnemers niet herkenbaar zijn

Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slap
Gemiddeld aantal publicaties per jaar

PSEUDONIMISEREN =

Data waarmee deelnemers niet herkenbaar zijn

CODE
Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slap
Gemiddeld aantal publicaties per jaar



Data waarmee deelnemers herkenbaar zijn (SLEUTELbestand):

CODE
Naam
E-mail
Werkgever + functie



Volledige dataset

Naam
E-mail
Werkgever + functie
Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slap
Gemiddeld aantal publicaties per jaar

Bewaard in beveiligd bestand op OneDrive
Mag gedeeld worden

SLEUTELBESTAND
Bewaard in AFZONDERLIJK beveiligd bestand op OneDrive
Mag NIET gedeeld worden

Re-identificatie is mogelijk

ANONIMISEREN =

Data waarmee deelnemers niet herkenbaar zijn

CODE
Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slap
Gemiddeld aantal publicaties per jaar

Data waarmee deelnemers herkenbaar



zijn:
CODE

Volledige dataset

Gender
Nationaliteit
Gemiddeld aantal stappen per dag
Gemiddeld aantal uren slap
Gemiddeld aantal publicaties per jaar

Bewaard in beveiligd bestand op OneDrive
Mag gedeeld worden

DEFINITIEF VERWIJDERD
(nergens nog een kopie bewaard)

Je behoudt enkel een anonieme dataset