

## Richtlijn Databeheerplan

### Achtergrondinformatie en uitleg Databeheerplan onderzoek Dimence Groep

#### Inleiding

De huidige privacywetgeving dwingt organisaties om op een verantwoorde manier met persoonsgegevens om te gaan. Op 2 mei 2016 is de Europese Algemene Verordening Gegevensbescherming in werking getreden. Dit heeft gevolgen voor wetenschappelijk onderzoek waarbij gewerkt wordt met persoonsgegevens. Vanaf mei 2018 is de verordening volledig van kracht en moet de bedrijfsvoering bij instellingen zó zijn ingericht dat men zich aantoonbaar conformeert aan de nieuwe bepalingen.

Van iedere onderzoeker van de Dimence Groep (DG) wordt verwacht dat hij of zij de volledige set van gebruikte onderzoeksdata "duurzaam" archiveert.

De DG wil er met deze richtlijn voor zorgen dat de kwaliteit, beschikbaarheid en toegankelijkheid van onderzoeksdata binnen de DG geborgd is en getoetst kan worden of voldaan wordt aan wet- en regelgeving en gedragscodes. De bij wetenschappelijk onderzoek gebruikte data moeten dus te allen tijde opvraagbaar en controleerbaar zijn.

De nieuwe richtlijn vloeit voort uit de Researchcode DG (zie [website CWO](#)), een gedragsvoorschrift waarin de DG duidelijker het belang van wetenschappelijke integriteit wilde onderstrepen.

Daarnaast is het doel van de richtlijn om onderzoekers zich meer bewust te laten zijn van de persoonlijke verantwoordelijkheid die ze hebben en hoe ze deze verantwoordelijkheid kunnen waarborgen.

Het is relevant om deze gegevens adequaat op te slaan omdat er binnen de DG veel (wetenschappelijk) onderzoek wordt uitgevoerd, waarbij het verzamelen, beheren, bewerken en analyseren van data een belangrijk onderdeel is. In de meeste gevallen gaat het om vertrouwelijke gegevens van deelnemers aan het onderzoek, waardoor het in het kader van privacy van belang is om op een goede en verantwoorde wijze om te gaan met de data. Bovendien moeten onderzoek en resultaten door derden gerepliceerd of geverifieerd kunnen worden. Daarnaast helpt zo'n plan de onderzoeker om ordening te houden in het proces van dataverzameling, analyse, en publicatie. Bij het opstellen van het databeheerplan voor de DG is gebruik gemaakt van de vuistregels van het Landelijk Coördinatiepunt Research Data Management (LCRDM<sup>1</sup>), die nog in ontwikkeling zijn.

Een databeheerplan is een één tot twee pagina's tellend document dat beschrijft hoe onderzoeksgegevens moeten worden verzameld, beheerd en gedeeld. Het databeheerplan is een integraal onderdeel van het onderzoeksproces en bij de DG opgenomen in het Aanmeldformulier patiëntgebonden onderzoek. Het doel van het databeheerplan is om vooraf in kaart te brengen welke data verzameld worden. Bovendien wordt beschreven waar data bewaard worden en wie toegang heeft tot de data.

---

<sup>1</sup> De juridische werkgroep van het Landelijk Coördinatiepunt Research Data Management (LCRDM) heeft tot taak de universitaire instellingen en het bredere onderwijsveld in Nederland beleidsmatig en praktisch te ondersteunen bij het borgen van de volgende aspecten van onderzoek: zeggenschap, intellectueel eigendom, gegevensbescherming en privacy. Het LCRDM is ingericht op verzoek van de VSNU en wordt gefaciliteerd door SURF.<sup>1</sup>

Door het databeheerplan:

- Zorgt de onderzoeker ervoor dat de onderzoeksgegevens aantoonbaar juist, volledig en betrouwbaar zijn door vooraf goed na te denken over welke gegevens nodig zijn en hoe de gegevens verzameld, geanalyseerd en opgeslagen worden.
- Verhoogt de efficiëntie van onderzoeksmethodes en voorkomt dubbel werk.
- Verbetert de beveiliging van gegevens en minimaliseert het risico van dataverlies.
- Verhoogt de replicerbaarheid van het onderzoek.

Het onderzoeksprotocol met aanmeldformulier (waarin het databeheerplan is opgenomen) wordt ingediend bij de Commissie Wetenschappelijk Onderzoek voor beoordeling (zie [www.dimencegroep.nl/onderzoek](http://www.dimencegroep.nl/onderzoek)). Daarnaast vindt beoordeling door de afdeling I&A plaats als er in het plan om bestaande data uit het EPD wordt verzocht. I&A bekijkt dan of de gevraagde data uit het systeem te halen is.

## **Bewerken, opslag en archivering data wetenschappelijk onderzoek**

### **Wat moet waar worden opgeslagen**

Alle (ruwe) onderzoeksdata moet worden opgeslagen; Van vragenlijsten tot video-opnames tot de stappen die genomen zijn om data te bewerken (bv in de vorm van een logboek). Al deze gegevens worden minstens tien jaar bewaard. Belangrijk is ook dat alle data compleet is, zodat "verificatie van het onderzoek" mogelijk is (repliceren onderzoek elders, verificatie resultaten door derden etc.)

Voor elke onderzoeksaanvraag die bij de Commissie Wetenschappelijk Onderzoek wordt ingediend wordt op de G-schijf van de DG een map op naam van de onderzoeker aangemaakt en ingericht. Hierin worden alle relevante documenten, zoals de aanvraag, het databeheerplan, verslag, syntax en dergelijke opgeslagen (zie 'Mappenstructuur G-schijf'). Ook onderzoeksdata, zoals verstrekte data en bewerkte data worden daarin geplaatst. De onderzoeker heeft toegang tot zijn/haar onderzoek en heeft de beschikking over alle documenten/data behorend bij dat onderzoek. Onderzoeksdata opgevraagd bij I&A wordt direct in deze omgeving geplaatst. Voor het veilig thuis werken op de G-schijf kan een token worden aangevraagd via de CWO.

### **Waar wordt al die data opgeslagen?**

#### Opslag hardcopy bestanden

- Alle hardcopy gegevens (vragenlijsten, formulieren, cd-roms, usb-sticks e.d.) worden tijdens de duur van het onderzoek door de onderzoeker bewaard in een afsluitbare kast.
- De documenten worden geanonimiseerd of gecodeerd bewaard.
- Nadat het onderzoek is afgerond worden de hardcopy gegevens in een doos (of dozen) bewaard in het archief van de Commissie Wetenschappelijk Onderzoek (CWO) aan de burg. Roelenweg in Zwolle. Op de doos (of dozen) wordt het volgende vermeld: naam van het onderzoek, voor- en achternaam onderzoeker, functie en werkplek, telefoonnummer, emailadres, stopdatum onderzoek en 'bewaren tot' datum. De gearchiveerde gegevens worden minimaal 15 jaar bewaard.
- Voor opslag van onderzoeksgegevens kan contact worden opgenomen met de CWO: [onderzoek@dimencegroep.nl](mailto:onderzoek@dimencegroep.nl). Onderzoekers zijn zelf verantwoordelijk voor het brengen van de gegevens en stemmen af met de CWO.

- Bestanden met data en databewerkingen
- Zowel de ruwe, opgeschoonde als de bewerkte bestanden (ook de oudere versies van bestanden) dienen bewaard te worden in de daarvoor bestemde omgeving op de G-schijf.
- Na goedkeuring door de RvB van het onderzoeksprotocol, krijgt de onderzoeker van de CWO toegang tot een eigen map op de G-schijf met bijbehorende mappenstructuur voor de dataopslag.
- In deze omgeving worden via een digitaal logboek de bewerkingen van databestanden en andere belangrijke beslismomenten bijgehouden en onderbouwd door de onderzoeker.
- Alle syntaxen (bv SPSS) van de data-analyses ten behoeve van databewerking en publicaties worden bewaard in dezelfde omgeving.
- Per publicatie wordt in een aparte map aangemaakt op de G-schijf: syntax, logboek, bronbestand, en alle versies van de bewerkte bestanden (inclusief back-up). Het proces van dataverwerking wat uiteindelijk leidt tot publicatie moet gereproduceerd kunnen worden.
- Onderzoekers mogen buiten de Dimence Groep omgeving alleen met gecodeerde en niet-tot-de-persoon-herleidbare onderzoeksgegevens werken.

#### Back-up gegevensbestanden

- Afdeling I&A draagt zorg voor het veiligstellen van alle gegevensbestanden op de G-schijf. Bij een calamiteit is het mogelijk de situatie, tot kort voor de calamiteit, te herstellen. De herstelwerkzaamheden worden uitgevoerd door de afdeling I&A.
- Per abuis verwijderde bestanden kunnen door I&A worden hersteld indien de aanvraag tot herstel geschied maximaal 4 weken na de verwijdering.

### **Mappenstructuur G-schijf**

De G-schijf bevat voor elk onderzoek een map met daarin submappen:

- Actuele versie onderzoeksprotocol
- Logboek
- Correspondentie
- Ruwe data
- Syntaxen
- Output na analyse
- Actueel manuscript
- Codering/sleutel

### **Wie is verantwoordelijk voor dataverzameling, analyse en archivering en voor gebruik van complete en correcte data?**

Primaire verantwoordelijkheid voor dataverzameling, analyse en archivering en voor gebruik van complete en correcte data ligt bij de onderzoeker.

Begeleiders (promotoren, onderzoeksbegeleiders, supervisors) hebben de taak onderzoekers te wijzen op deze verantwoordelijkheid.

Verlies van onderzoeksdata of andere datalekken in het kader van onderzoek moeten gemeld worden bij de staffunctionaris Privacy en Informatieveiligheid van de DG en de CWO (zie ook: Procedure melden potentieel datalek).