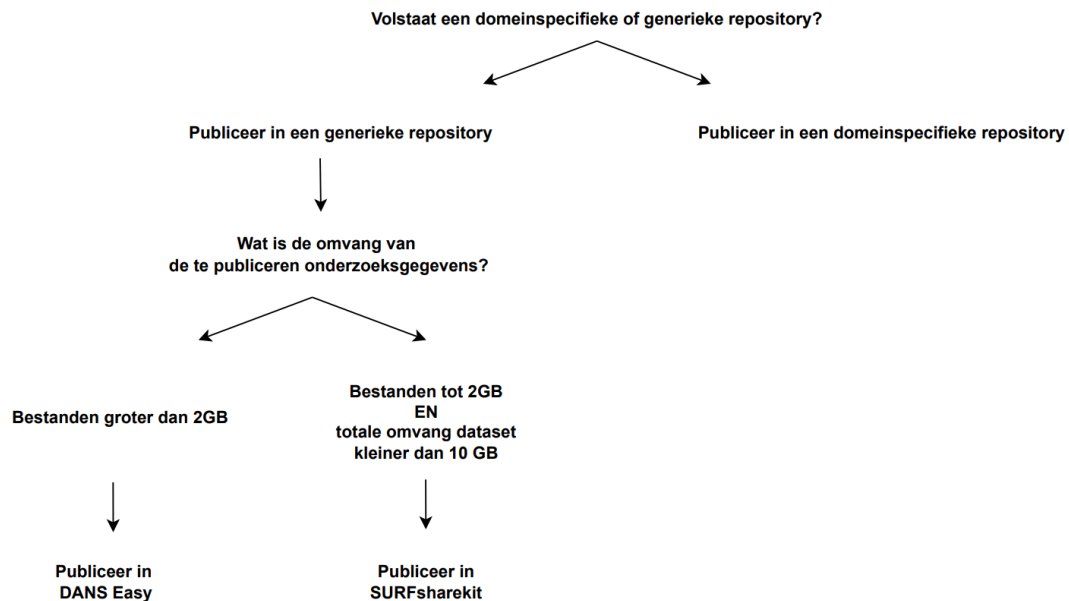


Proces publicatie onderzoeksgegevens

Een onderzoeker van hogeschool Leiden kan voor het publiceren van onderzoeksgegevens gebruik maken van een van de volgende data-repositories:

- Publicatie in de DANS Easy repository
- Publicatie in de HBO-kennisbank (d.m.v. SURFsharekit)
- Publicatie in een domeinspecifieke repository

Elk van deze repositories heeft eigen eisen op het gebied van bestandsformaten, datadocumentatie en metadatering. In dit document wordt het proces beschreven wat wordt gevolgd wanneer een onderzoeker onderzoeksgegevens wilt publiceren in een van de genoemde oplossingen. Het onderstaande schema toont welke afwegingen worden gemaakt om een repository te selecteren:



Proces publicatie in DANS Easy

Stap 1: Een onderzoeker heeft in het datamanagementplan aangegeven of geeft bij een onderzoekondersteuner van hogeschool Leiden aan onderzoeksgegevens te willen publiceren in de DANS Easy repository.

Vereiste voor het verdere proces:

- Het datamanagementplan is ingevuld
- Het CGA-formulier is ingevuld

Stap 2: De datasteward (centraal/decentraal) plant een afspraak in met de onderzoeker om het datamanagementplan en het proces voor datapublicatie te bespreken. Voor publicatie in DANS Easy is een formulier voor **metadata*** en een richtlijn voor **datadocumentatie (REAMDE-file)*** beschikbaar.

Stap 3: In het gesprek tussen de datasteward (centraal/decentraal) en de onderzoeker worden de volgende onderwerpen besproken:

- a. Publicatievoorwaarden onderzoeksgegevens
 - Zijn er afspraken gemaakt over het eigenaarschap van de te publiceren (onderzoeks)gegevens?
 - Data van externen, samenwerkingsverbanden, etc.
 - Controle van het Informed Consent
 - Is er een Informed Consent formulier dat archivering en publicatie van onderzoeksgegevens mogelijk maakt?
- b. Selectie van persoonsgegevens geschikt voor publicatie
 - Bevat de data(set) persoonsgegevens?
 - Vaststellen van de publicatievorm (Open of Restricted Access)
Hierbij wordt gebruik gemaakt van de LCRDM richtlijn: Risicomanagement voor data over mensen.
- c. Metadata en datadocumentatie
 - Metadata
 - De datasteward en de onderzoeker bespreken de metadata die benodigd is bij de DANS Easy repository Hierbij wordt gebruik gemaakt van het **metadata*** formulier.
 - Datadocumentatie
 - Daarnaast wordt de richtlijn voor **datadocumentatie*** besproken, vervolgens wordt de onderzoeker gevraagd om een eerste opzet te maken voor een README file.
- d. Bespreken van de vervolgstappen voor het verdere publicatieproces.

Stap 4: De onderzoeker wordt gevraagd om op Research Drive/ Office Groups een Datapackage (te publiceren gegevens + README file) klaar te zetten. De datasteward krijgt leesrechten op de data package en voert de volgende controle uit:

- a. De README file (**datadocumentatie***) wordt vergeleken met de inhoud van de dataset en indien nodig worden benodigde aanpassingen gemaakt.
- b. Controle of de databestanden zijn omgezet naar de lijst van duurzame bestandsformaten van DANS Easy.

Stap 5: Na afronding van het project deponeert de datasteward de gegevens in DANS Easy. De onderzoeker ontvangt vervolgens een bevestiging van de publicatie met de DOI.

Proces: Publicatie in de HBO-kennisbank (d.m.v. SURFSharekit)

Stap 1: Een onderzoeker heeft in het datamanagementplan aangegeven of geeft bij een onderzoekondersteuner van hogeschool Leiden aan onderzoeksgegevens te willen publiceren in de HBO kennisbank.

Vereiste voor het verdere proces:

- Het datamanagementplan is ingevuld
- Het CGA-formulier is ingevuld

Stap 2: De datasteward (centraal/decentraal) plant een afspraak in met de onderzoeker om het datamanagementplan en het proces voor datapublicatie te bespreken. Voor publicatie in de HBO-kennisbank is een richtlijn voor **datadocumentatie (README-file)*** beschikbaar.

Stap 3: In het gesprek tussen de datasteward (centraal/decentraal) en de onderzoeker worden de volgende onderwerpen besproken:

- a. Publicatievoorwaarden onderzoeksgegevens
 - Zijn er afspraken gemaakt over het eigenaarschap van de te publiceren (onderzoeks)gegevens?
 - Data van externen, samenwerkingsverbanden, etc.
 - Controle van het Informed Consent
 - Is er een Informed Consent formulier dat archivering en publicatie van onderzoeksgegevens mogelijk maakt?
- b. Selectie van persoonsgegevens geschikt voor publicatie
 - Bevat de data(set) persoonsgegevens?
 - Vaststellen van de publicatievorm (Open, Restricted of Closed Access)
Hierbij wordt gebruik gemaakt van de LCRDM richtlijn: Risicomanagement voor data over mensen.
- c. Metadata en datadocumentatie
 - Metadata
 - De metadata die benodigd is voor publicatie in de HBO-kennisbank komt grotendeels overeen met een standaard publicatie in de HBO-kennisbank. De onderzoeker wordt om deze reden gevraagd een samenvatting van het onderzoek aan te leveren en tags te selecteren.
 - Datadocumentatie
 - Daarnaast wordt de richtlijn voor **datadocumentatie*** besproken, vervolgens wordt de onderzoeker gevraagd om een eerste opzet te maken voor een README file.
- d. Bespreken van de vervolgstappen voor het verdere publicatieproces

Stap 4: De onderzoeker wordt gevraagd om op Research Drive/ Office Groups een Datapackage (te publiceren gegevens + README file) klaar te zetten. De datasteward krijgt leesrechten op de data package en voert de volgende controle uit:

- a. De README file (**datadocumentatie***) wordt vergeleken met de inhoud van de dataset en indien nodig worden benodigde aanpassingen doorgegeven.

- b. Controle of de databestanden voldoen aan de maximale bestandsgrote van 2GB en of de totale data package kleiner is dan 10GB.

Stap 5: De informatiespecialisten van het Mediacentrum publiceren de dataset in de HBO-kennisbank (d.m.v. SURFSharekit). Voor het verzenden van de onderzoeksgegevens naar het Mediacentrum wordt gebruik gemaakt van SURFFilesender.

Stap 6: De onderzoeker ontvangt een bevestiging van de publicatie met de DOI.

Proces: publicatie in een domeinspecifieke repository

De in dit document beschreven oplossingen (publicatie in DANS Easy en publicatie in SURFSharekit) zijn beide generieke data-repositories. Deze repositories ondersteunen veel verschillende soorten onderzoeksgegevens en schrijven veelal generieke metadataschema's voor. Hierdoor zijn deze repositories voor veel onderzoeken geschikt. Het is echter mogelijk dat deze repositories voor sommige onderzoeksprojecten niet volstaan, in deze gevallen kan gebruik worden gemaakt van een domeinspecifieke repository. Voor het selecteren van een domeinspecifieke repository kan gebruik worden gemaakt van Re3data.org, dit is een wereldwijd register van data-repositories.

Bij publicatie in een domeinspecifieke repository wordt (deels) afgeweken van het in dit document opgestelde proces. In deze gevallen zullen de datastewards samen met de onderzoeker kijken naar de specifieke eisen van de repository.

Opmerkingen

*De formulieren: **metadata** en **datadocumentatie** zijn deliverables die tijdens de pilot zijn gerealiseerd.

** De processen die in dit document zijn beschreven zijn gebaseerd op een eerdere pilot van De Haagse Hogeschool.