

De Haagse Hogeschool op weg naar Open Science: van Research Drive naar DANS Easy

Rapport van geleerde lessen

Projectbeschrijving

Met deze pilot werd kennis opgedaan voor de ondersteuning van de laatste fasen van onderzoek, te weten a) publiceren en b) archiveren van data. Op De Haagse Hogeschool is DANS Easy het aanbevolen data-archief. Het deponeren in DANS Easy wordt daarom actief ondersteund door het team datastewards. In deze pilot werd een interne procedure gedefinieerd en tools (formulieren) ontwikkeld om efficiënte en uniforme ondersteuning te bieden aan alle onderzoekers werkzaam op de Haagse Hogeschool. Volgens deze procedure adviseren datastewards de onderzoekers over het voorbereiden van dataset en de keuze van data die gedeponeerd moeten worden. Vervolgens controleren datastewards de dataset en voeren de dataset in DANS Easy op basis van een metadataformulier aangeleverd door onderzoeker.

Voor de pilot zijn vier bestaande datasets gekozen van afgeronde onderzoekprojecten. De datasets zijn gecontroleerd voor de opname in DANS Easy. Drie van de datasets werden goedgekeurd en voorbereid voor de opname.

Bij De Haagse Hogeschool wordt momenteel de Research Drive aanbevolen als veilige opslag tijdens het onderzoekproject. Daarom werd in deze pilot uitgezocht hoe de veilige transfer van Research Drive naar Dans Easy geïntegreerd kan worden met de interne deponeerprocedure.

Resultaten

1. Beschrijving van een workflow: een voorstel voor het inrichten van interne procedure die ingezet kan worden bij het deponeren van alle datasets geproduceerd op De Haagse Hogeschool. Deze workflow kan ook gebruikt worden wanneer de data in een ander data-archief gedeponeerd zouden worden.
2. Aanmeldformulier voor onderzoekers: metadata en Read.me file instructie. De metadata velden zijn grotendeels overeenkomstig met Dublin Core Metadata gehanteerd door DANS Easy. Onderzoekers worden gevraagd alleen zelf de metadata op te geven, die de datasteward niet zelf kan invullen op basis van aangeleverd dataset. Het Read.me-file-formulier is een instructie voor het schrijven van aanvullende informatie over de dataset die toegevoegde waarde voor de toekomstige gebruikers van de dataset heeft. Deze formulieren zijn dus specifiek bedoeld voor het deponeren van data in DANS Easy.
3. Controleformulier van datasets voor datastewards. Dit controleformulier kan gebruikt worden voor het toetsen van alle datasets aan de eisen van de hogeschool, het archief en FAIR, ongeacht het archief waarin de dataset gedeponeerd wordt.
4. Rapport van geleerde lessen
5. Beleidsadvies: de samenhang tussen RDM en de fasen van het onderzoek

Geleerde lessen

Wat moet je overwegen bij de keuze tussen DANS Easy en Data Verse?

De Haagse Hogeschool heeft gekozen voor DANS Easy boven Data Verse vanwege de kosten verbonden aan de licentie voor Data Verse. Datasets die ontstaan op de Haagse Hogeschool zijn klein genoeg om het publiceren in DANS Easy kosteloos te houden. Een nadeel van deze keuze is dat wanneer wij op later termijn Data Verse zouden willen gebruiken, de datasets die eerder gepubliceerd waren in DANS Easy niet kunnen worden overgebracht naar Data Verse.

Hoe sluit je de samenwerking tussen je instelling en DANS Easy af?

De Haagse Hogeschool heeft een DPIA toets gedaan voor het proces van archiveren van data in DANS Easy. Op basis van deze toets heeft de Privacy Officer van de Haagse Hogeschool geconstateerd dat een DPIA niet nodig was. Een datasteward heeft een startgesprek gevoerd met contactpersoon van DANS Easy. Om een institutionele samenwerking met DANS Easy te beginnen is het nodig om een deponeren en verwerkerovereenkomst te tekenen. Deze documenten werden geconsulteerd met de privacy officer van de Haagse Hogeschool.

Welke soorten data kunnen openbaar gepubliceerd worden?

De ruwe versies van interviews kunnen nooit openbaar gepubliceerd worden. De interviews kunnen nooit volledig geanonimiseerd worden.

Video en audio mogen openbaar gepubliceerd worden, maar niet wanneer mensen in de opnames voorkomen.

Codeboeken en Read.me file kunnen altijd openbaar gepubliceerd worden. Mits ze juist zijn opgesteld en dus geen persoonsgegevens of privacygevoelige informatie bevatten.

Hoe werkt het publiceren onder embargo periode bij DANS Easy?

Wanneer de dataset onder embargo gepubliceerd wordt in DANS Easy, betekent het dat de dataset zelf niet toegankelijk is, maar ook dat de metadata van dataset niet zichtbaar zijn in de repository.

Waarom kan een dataset afgekeurd worden voor publicatie in DANS Easy?

Een van de datasets die deel uitmaakte van de pilot werd niet opgenomen in DANS Easy. De volgende reden speelde hier een rol: het onderzoek werd gedaan in consortiumverband, maar het eigenaarschap van de data was niet vastgesteld met de partners. Ook werden de deelnemers aan het onderzoek niet geïnformeerd dat de data gedeponereerd zouden worden in DANS Easy.

Wat zijn twijfels die spelen bij onderzoekers over het deponeren van data?

Vanuit gesprekken met onderzoekers zijn er verschillende twijfels over het publiceren van data naar voren gekomen. De onderzoekers zijn bang dat hun data op een verkeerde manier gebruikt of geïnterpreteerd zullen worden. Het archiveren of publiceren van data ook als ze anoniem zijn wordt beschouwd als een drempel omdat je hiervoor toestemming moet vragen aan de deelnemers van het onderzoek.

Waarom is het belangrijk om het eigenaarschap van data af te spreken binnen het lectoraat?

Wanneer de data gedeponereerd wordt onder "restricted access" moet een persoon aangewezen worden die toestemming zal geven voor het delen van data op verzoek. Dit kan een onderzoeker zijn die de data

heeft verzameld. Het kan echter gebeuren dat de onderzoeker de instelling verlaat. Daarom is het verstandig om de regie over data te leggen bij een lector (functie en niet een persoon).

Wanneer wordt een DOI toegekend aan een dataset door DANS Easy?

Een DOI kan “gereserveerd” worden aan het begin van de invoerprocedure. Dit betekent dat de dataset kan geciteerd worden al voordat het gepubliceerd is. Het is mogelijk om de invoerprocedure te onderbreken en het item opslaan voor later. Op die manier wordt de al gereserveerde DOI link bewaard.

Wat kunnen onderzoekers doen om het publiceren van data efficiënter te maken?

De belangrijkste conclusie van de pilot is: het publiceren van data is moeilijk of soms zelfs onmogelijk zonder bewustzijn van en keuzes door onderzoekers op het gebied van Research Data Management in de beginfase van een onderzoeksproject, en de verschillende acties

Hoe kan het beleid helpen om de kwaliteit van datasets te vergroten?

Om de procedure van archiveren makkelijker en efficiënter te maken voor de onderzoekers en datastewards is het belangrijk om beleid te laten vaststellen, voornamelijk het schrijven van een Data Management Plan (DMP) verplicht te maken voor alle onderzoeksprojecten (ook zonder voorwaarden van subsidieverstrekkers die het publiceren van data vereisen).

Wat kunnen datastewards doen om de kwaliteit van datasets te vergroten?

Het is verstandig dat datastewards tijdens onderzoeksprojecten tussengesprekken voeren met de onderzoekers om te controleren of de onderzoekers hun datamanagementplan naleven. Bovendien zou het voor datastewards inzichtelijk moeten zijn in welke fases de verschillende onderzoeksprojecten zich bevinden, zodat de RDM-ondersteuning goed kan worden toegesneden op de praktijk van het onderzoeksproject. Hiervoor zou een monitoringstool behulpzaam kunnen zijn, bijvoorbeeld een onderzoeksinformatiesysteem (CRIS). Niettemin blijft het toepassen van het DMP de verantwoordelijkheid van onderzoeker.

Hoe kan de data veilig overgedragen worden van Research Drive naar DANS Easy?

De onderzoeker voert dataselectie uit en deelt de opgeschoonde dataset met de datasteward via Research Drive. De datasteward krijgt alleen de rechten om de dataset in Research Drive te kunnen inzien. Hij kan de data niet verwijderen, verder delen of veranderingen aanbrengen in de dataset. De datasteward controleert de dataset en communiceert met onderzoeker eventuele punten ter verbetering. De datasteward gebruikt de koppeling van Research Drive met de verkenner om de dataset veilig te uploaden in DANS Easy waardoor het niet nodig is om de dataset eerst te moeten downloaden van Research Drive om de dataset te kunnen uploaden in DANS Easy.