

Pilot: Open science made simple: Ecosysteem en werkwijze voor hele onderzoekscyclus

Hogeschool: HZ University of Applied Sciences

Lectoraten:

- HZ lectoraat Data Science
- HZ lectoraat Onderwijsexcellentie

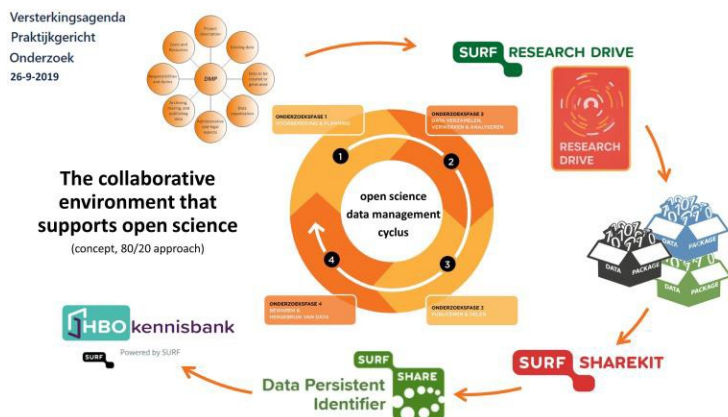
Contactpersoon: Kees Bal | kees.bal@hz.nl

Looptijd: februari 2021 – augustus 2021



Beschrijving van de pilot

Op de HZ is in een pilot een compact open-science ecosysteem en bijbehorende werkwijze gemaakt en getest. Dit is gebeurd in samenhang met enkele uitbreidingen in SURFsharekit. De uitkomst is een simpele 'plug and play' RDM aanpak voor onderzoeksprojecten.



Het doel van de pilot

De pilot beoogt onderzoekers (in eerste instantie, maar niet exclusief, hbo lectoraten) een eenvoudige oplossing te bieden om gedurende de hele onderzoekscyclus goed met hun data om te gaan.

Welk vraagstuk of probleem heb je aangepakt?

Het beleid voor RDM hebben veel hogescholen, waaronder de HZ, reeds op papier. Echter op de logische vraag van de onderzoekers: "Hoe dan?" (welke tools/software voldoen, en hoe toe te passen?), was het antwoord nog niet makkelijk te geven. Met name voor wat kleinere organisaties, die nog niet ver zijn en geen grote ondersteunende afdeling hebben, is dit een uitdaging.

De in dit project gevonden en getoetste werkwijze kan voor veel onderzoeksgroepen/- projecten dat antwoord zijn. De resultaten zijn een aanvulling op het bestaande aanbod, al naar gelang de context (aard project, voorkeur onderzoeker, specifieke hogeschool) kan de meest passende aanpak gekozen worden.

Samengevat: door het project uiteen te splitsen in diverse volgtijdelijke activiteiten en deze onder het projectteam te verdelen wist iedereen wat hij/zij wanneer moest bijdragen. Dit is gecombineerd met maandelijkse voortgangsmetingen waardoor we tussentijds konden afstemmen en elkaar konden helpen. Op een enkele verschuiving na, hebben we op deze manier de projectplanning kunnen volgen zoals vooraf bedacht.

Resultaten & impact

Resultaten

Het doel is behaald: er is een simpele werkwijze gevonden die onze onderzoekers, gedurende de hele onderzoekscyclus concrete handvatten biedt om conform het HZ datamanagementbeleid te werken. Deze werkwijze is tevens getest door twee onderzoekers van de HZ en in een validatiesessie besproken met onderzoekers en bibliotheekmedewerkers van de Haagse Hogeschool. Zie de presentatie (pdf) voor de twee casussen.

Gerealiseerd dan wel benoemd zijn:

- Procesgang en verantwoordelijkheden op de HZ betreffende onderzoeksdata vanaf start onderzoek (DMP), via Research Drive tijdens onderzoek, tot aan SURFsharekit voor publicatie;
- Uitbreiding SURFsharekit repository waardoor ook kleine tot middelgrote datasets inclusief metadata geüpload kunnen worden;
- Tips/instructies voor het maken van een readme voor de data-packages;
- Toevoeging van DOI binnen de SURFsharekit omgeving.

Datasets op HBO Kennisbank:

- https://www.hbo-kennisbank.nl/details/sharekit_hz_dataset:oai:surfsharekit.nl:a6b47c0c-91f6-425f-b5b8-9bfe7103de26
- https://www.hbo-kennisbank.nl/details/sharekit_hz_dataset:oai:surfsharekit.nl:46b6a8fe-9e9e-46aa-b636-52a7ea9ca952

Meer specifiek over de ervaringen en resultaten:

Datamanagementplan

Het bestaande datamanagementplan voor de HZ is aangepast met de input van de betrokken onderzoekers bij het project. Vooral op de punten van privacy, licenties, kwaliteit van de data en persoonsgegevens bleek het oude DMP niet voldoende bruikbaar. De plek waar het DMP aan bod komt in de workflow van Research Drive is naar voren gehaald.

Werkwijze Research Drive

De bestaande werkwijze voor Research op de HZ is in overleg met de ICT-afdeling aangepast op de punten die van belang zijn voor de eerstelijns ondersteuning. Er is nauwere betrokkenheid en inzet van IT bij Research Drive.

- Er is een opzet gemaakt voor de eerstelijns (HZ) en tweedelijns-ondersteuning (SURF) via het Topdesk-systeem van de HZ. Via een aparte tegel Research Drive kunnen projecten aangevraagd worden, het DMP via een invulformulier ingevuld worden en vragen/problemen over RD gesteld worden. Deze vragen worden afgehandeld door een van de behandelaars van de behandelaarsgroep (medewerkers mediatheek en ICT) van de HZ in Topdesk.
- Ook wordt de pagina verder aangevuld met informatie van SURF en FAQ's. Dit kunnen ook inhoudelijke vragen zijn over onderwerpen als README files of DMP's.
- In samenwerking met SURF is een template aangemaakt voor datasets, de bestaande basis van onderzoek (Edustandaard MODS) is uitgebreid met een aantal velden. Door SURF zijn aanvullende afspraken gemaakt met DANS Easy over de mogelijkheid van linken van grotere datasets.
- In SURFsharekit zijn 2 datasets ingevoerd door 2 verschillende onderzoekers, waarvan 1 bestaande uit een dataset en bijhorend artikel. Bij de losse dataset is een DOI toegevoegd. Bij de combinatie van een dataset en artikel in SURFsharekit zal nog verder gekeken moeten worden naar het linken vanuit het artikel naar de dataset en vice versa.
- Het was zoeken naar de plek van de README file, deze werden uiteindelijk als losse bestanden opgenomen.

Wat heb je geleerd?

Dat het inderdaad ook mogelijk is om data op te slaan en te publiceren via SURFsharekit. Het blijkt dat dit een gewaardeerde aanvulling is binnen het bestaande pallet aan datadiensten (dit is vooraf, gedurende en na de pilot gebleken op momenten dat de pilot gepresenteerd werd in bredere gremia). Ook hebben de betrokken onderzoekers aangegeven dat dit aansluit bij hun behoeften en dat ze er graag verder mee aan de slag gaan.

Hoe ga je hiermee verder?

De pilot is ontstaan vanuit een vraag van de HZ lectoren die geconfronteerd met het HZ datamanagement beleid vroegen om concrete handvatten en instructies. Aangezien de gevonden werkwijze succesvol blijkt zal deze werkwijze en de bijbehorende instructies als appendix aan het HZ datamanagementbeleid toegevoegd worden. Dit betekent dus dat deze werkwijze in ieder geval op de HZ de aangeraden route wordt.

Tegelijkertijd is het veel breder toepasbaar en zullen andere onderzoekers de nieuwe functionaliteiten ook kunnen gaan inpassen in hun aanpak. Wij willen hier als HZ graag behulpzaam in zijn en onze ervaringen en tips delen met deze partijen. Ook zullen we daarvoor materialen maken zoals bijvoorbeeld een informerend filmpje (dit wordt op dit moment gemaakt en zal binnen enkele weken nagestuurd worden). We nemen dus graag deel aan het vervolg - op initiatief van SURF - waarbij 10 Hogescholen betrokken zijn.

Wat is de impact op (het faciliteren) van praktijkgericht onderzoek?

Wij verwachten dat onderzoeksgroepen (met name in hogescholen waar nog niet alles ingericht is voor open-science en die minder data-intensieve onderzoeksprojecten doen) één op één de resultaten kunnen overnemen en ook deze werkwijze kunnen gaan implementeren. Met de uitkomsten van deze pilot worden dus praktijkgerichte onderzoekers ontzorgd en gefaciliteerd, en kan onderzoeksdata gemakkelijk gedeeld en gevonden worden (passend bij de open science ambities van het hbo).

Betrokken SURF-diensten

Welke SURF-diensten waren onderdeel van de pilot?

De pilot betrof primair SURFsharekit en secundair Research Drive.

Hoe zijn deze diensten ingezet of gebruikt binnen de pilot?

De functionaliteit van SURFsharekit is uitgebreid:

- Er is een nieuwe categorie van publicaties toegevoegd: de dataset;
- Daarbij behorend is een metadata invulpagina toegevoegd;
- Verder is ook de mogelijkheid tot het aanvragen van een DOI voor publicaties toegevoegd.

Betrokkenen en inzet

Welk lectoraat(en) of kenniscentra waren betrokken bij de pilot?

- HZ lectoraat Data Science (Loek van der Linde)
- HZ lectoraat Onderwijsexcellentie (Ingrid Snijders)

Welke inzet en expertise is vanuit SURF geleverd?

SURF heeft daadwerkelijk enkele productaanpassingen kunnen maken in SURFsharekit, en heeft in het proces meegedacht over hoe we het aanbod zo generiek en robuust mogelijk kunnen maken. Ook staat SURF centraal in het denken over een verdere uitrol van deze nieuwe faciliteiten. Idealiter staan deze spoedig ter beschikking van alle gebruikers van SURFsharekit, met daarbij ook een duiding van welke diensten in welke situaties aangeraden of mogelijk zijn.

Lessons learned en aandachtspunten

- De verschillen tussen de behoeften aan hulp en ondersteuning van onderzoekers zijn erg verschillend, het is van belang en leerzaam om hier samen in op te trekken en de behoeften duidelijk te krijgen.
- Van deze pilot hebben we geleerd dat een onderzoeker vraagstukken in een latere fase kan voorkomen door al in de DMP fase over vragen t.a.v. privacy, licenties, kwaliteit van de data en persoonsgegevens na te denken en de antwoorden te documenteren.
- De mediatheek van de HZ heeft tot nog toe geen ervaring met het invoeren van onderzoekspublicaties. Het ondersteunen bij en het publiceren van Open Access publicaties vraagt om grotere kennis van auteursrecht op onderzoekspublicaties en Open Access publiceren. Hier staan we dus nog aan het begin van de leercurve. Hopelijk kunnen we via het landelijke (DCC) netwerk hierin leren.
- Het leren maken van een README bijhorend bij een dataset is voor de meeste onderzoekers nieuw. Je dit zelf via de instructies aanleren zal de eerste paar keer iets meer tijd kosten.
- Zoals eerder vermeld, bij de combinatie van een dataset en artikel in SURFsharekit zal nog verder gekeken moeten worden naar het linken vanuit het artikel naar de dataset en vice versa.
- Tot slot, ook voor (hele grote) datasets die niet op SURFsharekit (kunnen) staan, is het mogelijk (en zelfs aangeraden) om een metadata veld aan te maken via SURFsharekit. De verwijzing naar de dataset vanuit SURFsharekit kan dan gewoon met een hyperlink. Zo vergroot je ook de vindbaarheid van deze data.