



Presentatie pilot op DCC Praktijkgericht Onderzoek-dag 11 oktober 2021
Kees Bal

DE PERSOONLIJKE HOGESCHOOL

14/10/21

Er was eens een mooi RDM beleid...

HZ Research Data Management (RDM) beleidskader concept april 2018



1. Achtergrond

In het Nationaal Plan Open Science (feb 2017) committeren de Nederlandse onderzoeksinstituten, waaronder de hbo-sector, zich aan het doel om in 2020 volgens de principes van *open science* te werken. Dat houdt in dat er voor publiek gefinancierd onderzoek *open access* (vrije toegang) is voor publicaties. Tevens betekent dit dat de (ruwe) onderzoeksgegevens volgens *open data* beschikbaar gemaakt worden. Subsidieerregelingen stellen op dit gebied ook steeds vaker specifieke eisen.

In de notitie 'Hogescholen en Open Science: Naar meer impact van praktijkgericht onderzoek' (VH, juni 2017) wordt specifiek gekeken naar het praktijkgerichte onderzoek dat door het hbo wordt verricht. Er staat: "Open Science in relatie tot praktijkgericht onderzoek betekent niet alleen openheid van publicaties maar betekent ook openheid over de hele breedte van producten van praktijkgericht onderzoek, zoals modellen, artefacten, protocollen, prototypen, demonstraties, etc. Deze openheid is wezenlijk voor het bevorderen van de circulatie en creatie van kennis, zodat de impact van praktijkgericht onderzoek wordt vergroot". De notitie onderstreept in deze tevens het belang van goed datamanagement.

Om bovenstaande nationale ambities waar te maken, en tegelijkertijd gevolg te geven aan de Algemene Verordening Gegevensbescherming (AVG 2018), is het van belang dat kennisinstellingen als HZ expliciet aandacht besteden aan verantwoorde omgang met onderzoeksgegevens. Onderhavig beleidskader geeft HZ onderzoekers, directie en bestuur hiervoor passende richtlijnen¹. Daarbij is getracht dit beleid zo hanteerbaar mogelijk te laten zijn.

Het moet er toe leiden dat onderzoeksdata die door een onderzoeker van HZ worden gegenereerd

- nauwkeurig, volledig, betrouwbaar en authentiek en voorzien van metadata;
- veilig met minimaal risico op verlies opgeslagen;
- traceerbaar en toegankelijk;
- en ontsluitbaar en citeerbaar

zijn volgens het FAIR principe (Findable, Accessible, Interoperable en Reusable).

2. Definities

Onder *onderzoeksdata* wordt verstaan: alle (digitale en niet digitale) gegevens die worden gegenereerd tijdens onderzoek, alsmede de instrumenten waarmee de onderzoeksdata zijn vastgelegd en andere relevante informatie (bijvoorbeeld vragenlijsten, software).

Onderzoeksdata omvatten:

- de primaire onderzoeksdata, de onbewerkte, ruwe metingen of observaties;
- de secundaire onderzoeksdata, data die ontstaan na bewerking door de onderzoeker (hercoderen, combineren, categoriseren, visualiseren, e.d.).

¹ Gebaseerd op het 'Universitair beleidskader onderzoeksdata Universiteit Utrecht' (2016)

Externe onderzoeksdata zijn onderzoeksdata die het onderzoek in licentie ter beschikking heeft gekregen om als bron te dienen, maar waarvan het eigenaarschap volledig buiten de HZ ligt (zoals statistische en financiële gegevens en survey data).

Te archiveren onderzoeksdata omvat alle onderzoeksdata uitgezonderd externe onderzoeksdata die relevant worden geacht om na afronding van het onderzoek te bewaren, bijvoorbeeld uit oogpunt van verificerbaarheid van het onderzoek en/of mogelijk hergebruik.

Onderzoeksdata hebben een nauwe relatie met:

- de onderzoeksobjecten zelf;
- de verslaggeving van het onderzoeksproces (bijv. in lab journaals);
- de onderzoeksresultaten, d.w.z. alle uitvindingen, uitkomsten, materialen, methodes, processen, producten (bijv. publicaties), programma's, software, vindingen of ontdekkingen die binnen een onderzoeksproject worden gegenereerd.

Onder *metadata* wordt verstaan: documentatie en/of informatie over de onderzoeksdata die nodig zijn om de inhoud en context van de data te begrijpen.

3. Doelen en bereik

- Voor onderzoekers van de HZ is duidelijk welke eisen aan de omgang met onderzoeksdata gesteld worden. Het beleid wil een kader stellen, zodat de kwaliteit, beschikbaarheid en toegankelijkheid van onderzoeksdata binnen HZ geborgd is en getoetst kan worden of voldaan wordt aan wet- en regelgeving en gedragscodes.
- Het beleid geeft duidelijkheid over de verantwoordelijkheden: wat de rol is van de verschillende actoren.
- Het beleid raadt de onderzoeker aan om met behulp van de beschikbare tools en middelen de onderzoeksdata beschikbaar te stellen van zijn onderzoek en/of afgeleverd te worden.

Het beleid is van toepassing op alle medewerkers van HZ, waaronder, maar niet beperkt tot:

- medewerkers van de faculteiten en de afdelingen;
- medewerkers van de afdelingen Gegevens/Algemene Verordening Gegevensbescherming; Informatie- en Mediatechnologie; en de afdeling Sociaalwetenschappelijke Onderzoek;
- Gedragscode praktijkgericht onderzoek voor het hbo.

Specifiek voor datamanagement gaat HZ additioneel uit van onderstaande uitgangspunten.

Creatie en verzameling

- Onderzoeksdata moeten worden verzameld, geregistreerd, geselecteerd en voor gebruik beschikbaar gesteld in overeenstemming met de van toepassing zijnde internationale normen.

² Deze tools zullen in de loop van 2018 in samenspraak met de onderzoekers vastgesteld worden.

Opslag, beheer en bewaring

- Het beheer van alle onderzoeksdata vindt plaats conform de binnen de HZ geldende vereisten rond informatiebeveiliging, privacy bescherming en transparantie (die in deze volgorde van belang gelden);
- Te archiveren onderzoeksdata moeten minimaal 10 jaar bewaard worden. De bewaartermijn gaat in op de datum van publicatie van de onderzoeksresultaten;
- De data worden veilig bewaard, beschermd tegen ongeautoriseerd gebruik en tegen verlies;
- Wie de rechthebbenden zijn van de data en wie gebruiksrechten hebben op de data is direct inzichtelijk;
- De te archiveren onderzoeksdata worden in een structuur bewaard geschikt voor langdurige opslag en latere raadpleging;
- Per afgerond onderzoeksproject wordt de volledige set van gebruikte onderzoeksdata voor zover relevant voor verificatie van het onderzoek;
- De onderzoeksdata zijn zodanig van metadata voorzien dat de data inzichtelijk beschreven worden en vindbaar zijn voor de rechthebbenden;
- Bewaartermijn en contactgegevens van de rechthebbenden worden vastgelegd;
- Eventuele vernietiging van de data wordt vastgelegd.

Vindbaarheid

- De data worden toegankelijk gemaakt en beschikbaar gesteld in een gebruiksvriendelijk mogelijke omgeving, op voorwaarde dat de data beschermd zijn tegen ongeautoriseerd gebruik;
- De data worden toegankelijk gemaakt en beschikbaar gesteld in een gebruiksvriendelijk mogelijke omgeving, op voorwaarde dat de data beschermd zijn tegen ongeautoriseerd gebruik;
- De te archiveren onderzoeksdata worden voor zover redelijkerwijs mogelijk en met inachtneming van gepaste voorzorgsmaatregelen beschikbaar gesteld voor toegang en hergebruik voor onderzoek binnen en buiten HZ. Exclusieve rechten voor hergebruik of publicatie van onderzoeksdata worden niet overgedragen aan commerciële uitgeverij en hun vertegenwoordigers zonder zelf het recht te behouden op het voor hergebruik open beschikbaar stellen van die data, tenzij dit een randvoorwaarde is voor verstrekking van fondsen.

5. Rollen en verantwoordelijkheden

HZ onderscheidt de volgende (groepen) actoren die elk eigen verantwoordelijkheden hebben op het gebied van datamanagement:

- Degenen die de data creëren: onderzoekers en studenten die participeren in HZ onderzoek. Bij hen ligt de primaire verantwoordelijkheid voor de zorg voor de onderzoeksdata;
- Directeuren. Zij hebben de verantwoordelijkheid om erop toe te zien dat hun studenten, onderzoekers en onderzoeksleiders zich houden aan het beleidskader. Indien nodig kunnen zij nadere richtlijnen vaststellen;

- College van Bestuur. Het College van Bestuur heeft de verantwoordelijkheid voor het (doen) opstellen van een algemeen beleidskader en voor het faciliteren van de uitvoering van het beleidskader, bijv. infrastructuur, voorlichting en andere vormen van ondersteuning.

Als algemeen uitgangspunt geldt dat iedere medewerker gehouden is om relevante data als een goed rentmeester te beheren.

Onderzoekers

Onderzoekers zijn gehouden om de verantwoordelijkheid van de individuele onderzoeker (of van de onderzoeksleider indien van onderzoekers betreft), om bij het begin van een onderzoeksproject een Data Management Plan (DMP) op te stellen en de afspraken in het DMP uit te voeren. In een data management plan wordt aangegeven op welke manier de onderzoeksdata verzameld, opgeslagen, beschermd, gedocumenteerd en gebruikt kunnen worden.

Onderzoekers zijn gehouden om:

- voor onderzoeksprojecten de verplichte rubrieken in het HZ 'Data Management Plan' template in te vullen;
- erfor te zorgen dat de integriteit, de vertrouwelijkheid en de veiligheid van de data gewaarborgd is;
- wanneer het gaat om data die vallen onder de Wet Bescherming Persoonsgegevens of onder de European Data Protection Guideline, zich bewust te zijn van hun verplichtingen en mogelijke wettelijke aansprakelijkheid;
- aan het einde van het onderzoeksproject of bij afloop van hun dienstverband aan de instelling te zorgen voor voortzetting van het operationeel beheer van de onderzoeksdata;
- in onderzoeksvoorstellen en subsidie-aanvragen rekening te houden met de kosten van het bewaren en beheren van onderzoeksdata.

Directeuren

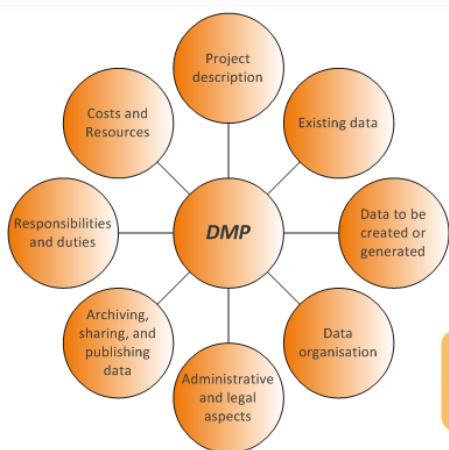
- zorgen ervoor dat onderzoekers zich bewust zijn van hun verantwoordelijkheden en verplichtingen ten aanzien van de onderzoeksdata die zij gedurende hun onderzoek verzamelen;
- faciliteren opleiding en training, daar waar zich hiaten in kennis en vaardigheden op het terrein van datamanagement voordoen.

College van Bestuur

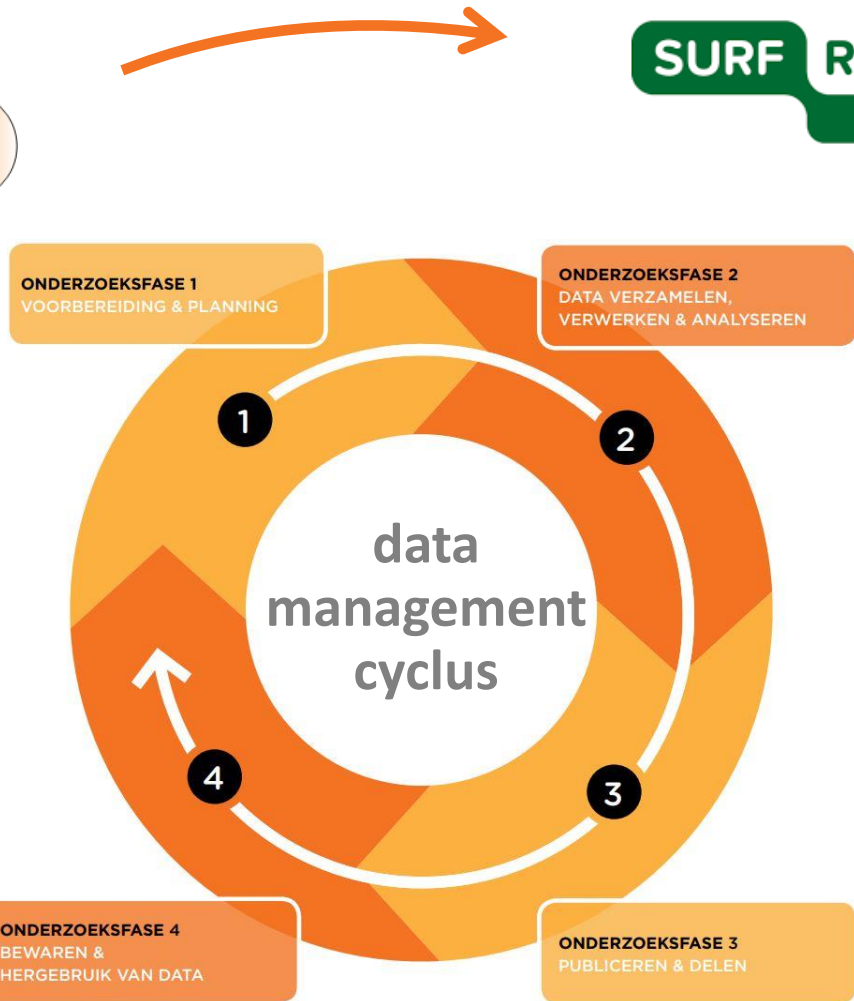
- is verantwoordelijk voor het doen opstellen en vaststellen van het RDM beleidskader;
- zorgt op instellingsniveau voor voorzieningen en diensten voor opslag en het veiligstellen van onderzoeksdata voor huidige en toekomstig gebruik. Dit betreft onder meer diensten voor opslag, backup, registratie, deponeren en duurzaam bewaren van dataverzamelingen en infrastructuur waarmee onderzoeksdata toegankelijk gemaakt kunnen worden en gedeeld kunnen worden met derden.

“Dat is prima, maar hoe doen we dat dan?”

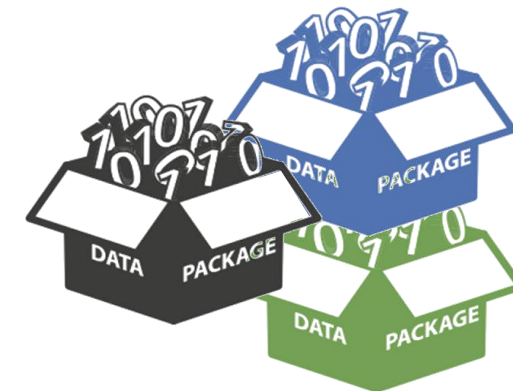
Idee?



A collaborative environment that supports open science



SURF RESEARCH DRIVE



SURF SHAREKIT



Data Persistent Identifier

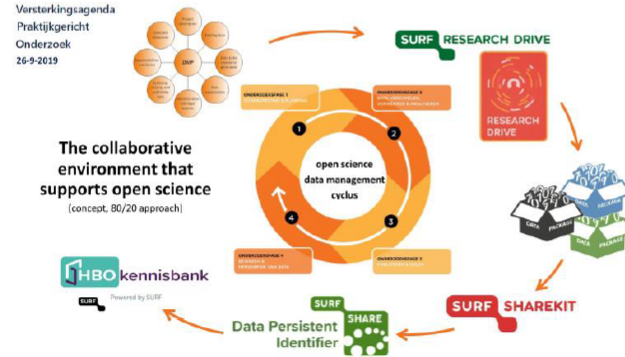


SURF Powered by SURF



Dat werd een voorstel voor het DCC:

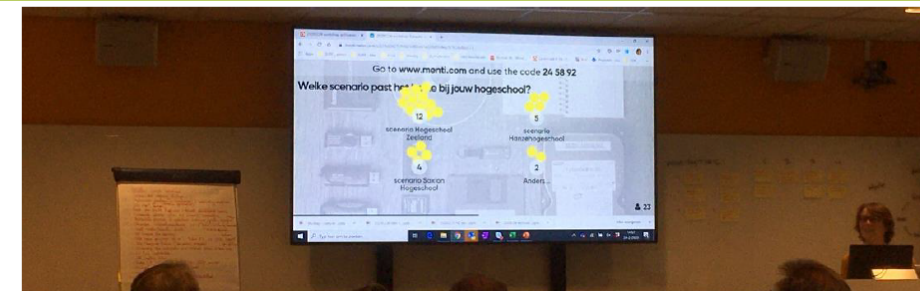
- *Geef een korte beschrijving van de pilot:* Op de HZ wordt in een pilot een compact open-science ecosysteem en bijbehorende werkwijze getest, geëvalueerd, geoptimaliseerd en geïmplementeerd, met bijbehorende documentatie. Een en ander in interactie met de reeds in gang gezette uitbreiding van SURF Sharekit. De beoogde uitkomst is een simpele doch effectieve 'plug and play' RDM aanpak voor onderzoeksprojecten.



- *Wat is 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 onderzoeksproces goed met hun data om te gaan.
- *Welk vraagstuk of probleem wil je aanpakken?:* 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?), is het antwoord 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 uitkomsten van dit project zullen voor veel onderzoeksgroepen/-projecten dat antwoord zijn. Deze simpele aanpak voor open-science is aanvullend op het bestaande. Al naar gelang de context (aard project, voorkeur onderzoeker, specifieke hogeschool) kan de meest passende aanpak gekozen worden.
- *Wat is de verandering of verbetering die met de pilot wordt beoogd?:* Een relatief simpele, integrale en niet commerciële oplossing (systeem plus werkwijze) voor RDM met open-science component, voor de hele onderzoeksproces van DMP tot publicatie, bestaat nog niet. Dit project voorziet hierin. De primaire doelgroep zijn onderzoeksgroepen in hogescholen waar nog niet alles ingericht is voor open-science en die minder data-intensieve (grenzen nader af te spreken met SURF) onderzoeksprojecten doen. Met de uitkomsten van deze pilot kunnen de groepen ontzorgd en gefaciliteerd worden.

Sectie 2 – Resultaat & impact

- *Welk resultaat wil je bereiken? Wat is er klaar aan het eind van de pilot?*
Er is een volledig werkend systeem (kernelementen zijn Research Drive als opslagomgeving en Sharekit als (data-)publicatieomgeving) een gedocumenteerde werkwijze. Anderen binnen en buiten de HZ kunnen hier direct mee aan de slag.
- *Wat is de beoogde impact op (het faciliteren) van praktijkgericht onderzoek?*
Dat hogescholen en al die onderzoeksgroepen, die nu worstelen met de vraag hoe ze eenvoudig doch zorgvuldig hun data-management kunnen doen en daarbij de principes van open-science willen toepassen, dit systeem en bijbehorende werkwijze kunnen adopteren. Uit eerste verkenningen (o.b.v. het figuur op pagina 3) binnen de groep deelnemers aan de Versterkingsagenda Praktijkgericht Onderzoek en op de SURF/DANS 'data-opslag dag' op 24 februari jl. waar 16 hogescholen aanwezig waren bleek hieraan een grote behoefte (zie foto hieronder, 12 stemmen!). NB: Het is niet de oplossing voor alles, met name zeer data intensieve projecten vragen een andere aanpak. Onze inschatting echter is dat voor 70% a 80% van de praktijkgerichte onderzoeksprojecten dit wel zou kunnen werken.
- *Hoe worden de resultaten van het pilot beschikbaar gesteld binnen de eigen hogeschool?*
De ambitie is te komen tot een systeem en werkwijze die expliciet als de voorkeursvariant in ons HZ beleid genoemd wordt. Dit impliceert een actieve promotie en ondersteuning vanuit de Dienst Onderwijs, Onderzoek en Kwaliteit. Dit is tevens de wens van onze lectoren.
- *In welke vorm wordt de opgedane kennis beschikbaar gesteld aan andere hogescholen?*
Aangezien vele hogescholen al kennis genomen hebben van dit idee (zie de 2^e vraag hierboven) is het zeker de bedoeling de resultaten breed en actief te verspreiden, één en ander in samenwerking met SURF. Indien het een succes is, verwacht ik dat er op vele conferenties interesse zal zijn. Sowieso worden de producten en documentatie publiek beschikbaar gemaakt.
- *Hoe kunnen andere hogescholen hun voordeel doen met de resultaten die in de pilot worden verkregen?:* Ook daar zullen onderzoeksgroepen voor dezelfde uitdaging staan. Omdat vele hogescholen reeds Research Drive en Sharekit gebruiken, verwacht ik dat ook zij vaak hiervoor zullen gaan kiezen.



Samen met SURF
is er gebouwd
en getest...

Met als resultaat:

Primeur!

- ✓ De grootste kennisbank van het HBO
- ✓ Inspiratie op jouw vakgebied
- ✓ Vrij toegankelijk

Je gekozen filters Wis filters

- Publicaties met bestand / URL ✕
- Dataset ✕

Publicatiejaar

Van Tot en met →

Type

- Afstudeerproduct (29)
- Bachelor (27)
- Master (2)
- Artikel (2)
- Rapport (1)
- Boekdeel (1)
- Dataset (1)

Hogeschool

- HZ University of Applied Sciences (1)

2 Resultaten

Sorteer op: Meest relevant

SCADA data windturbines in Nederland

L.W. van der Linde (Onderzoeker)

Een dataset bestaande uit SCADA data van drie windturbines. Daarbij toegevoegd data van de Dutch Offshore Wind Atlas zodat deze in verband gebracht kan worden met de SCADA data. Als laatst is ook een klein praktisch onderzoek opgenomen. Dat onderzoekt aan de hand van machine learning of er op basis van ...

2021-07-01 • Techniek • Dataset • Open Access

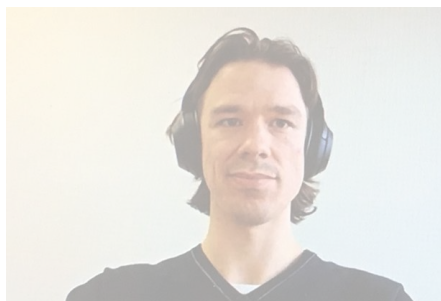
Data from: Relationship quality time: the validation of a relationship quality scale in higher education

I Snijders (Onderzoeker); L. Wijnia; R.M.J.P. Rikers; S.M.M. Loyens

Data on which the manuscript was accepted: Reference: Snijders, I., Rikers, R. M. M., Wijnia, L., and Loyens, S. M. M., "Relationship Quality Time: The validation of a relationship quality scale in higher education". Manuscript submitted for publication on 5 November 2016 at the journal of Higher Education ...

2017-11-22 • Opvoeding & Onderwijs • Dataset • Open Access

Loek van der Linde



SCADA data windturbines in Nederland

L.W. van der Linde (Onderzoeker)

Bestanden downloaden (4)

Rechten: CC BY-NC-SA

Samenvatting

Een dataset bestaande uit SCADA data van drie windturbines. Daarbij toegevoegd data van de Dutch Offshore Wind Atlas zodat deze in verband gebracht kan worden met de SCADA data. Als laatst is ook een klein praktisch onderzoek opgenomen. Dat onderzoekt aan de hand van machine learning of er op basis van de dataset kan worden geconcludeerd of er erosie aan de turbinebladen plaats heeft gevonden.

NB: de dataset is gevoelig voor concurrentie. Daarom is de dataset geanonimiseerd en zijn de data dusdanig aangepast dat deze niet naar de reële data en exploitant herleidbaar zijn.

Trefwoorden

scada

erosie

machine learning

crisp-dm

windmolens

windturbines

Thema

Techniek



Organisatie	HZ University of Applied Sciences
Afdeling	Domein Technology, Water & Environment
Lectoraat	Lectoraat Data Science
Datum	2021-07-01
Type	Dataset
DOI	10.48544/a6b47c0c-91f6-425f-b5b8-9bfe7103de26
Taal	Nederlands

Data from: Relationship quality time: the validation of a relationship quality scale in higher education

I Snijders (Onderzoeker); L. Wijnia; R.M.J.P. Rikers; S.M.M. Loyens

Bestanden downloaden (2)

Rechten: [Alle rechten voorbehouden](#)

Ingrid Snijders



Samenvatting

Data on which the manuscript was accepted:

Reference: Snijders, I., Rikers, R. M. M., Wijnia, L., and Loyens, S. M. M., “Relationship Quality Time: The validation of a relationship quality scale in higher education”. Manuscript submitted for publication on 5 November 2016 at the journal of Higher Education for Research and Development (HERD), and after revision accepted.

Period on which the data was collected: start of 11 May 2015 (3 June 2015 reminder) till 24 June 2015

The files contain confidential (quantitative and qualitative) data, e.g., student numbers, names, and, information that can be traced down to individual persons such as by age, gender, study program, or, explicit



Organisatie	HZ University of Applied Sciences
Afdeling	Domein Health, Education & Wellbeing
Lectoraat	Lectoraat Onderwijsexcellentie
Datum	2017-11-22
Type	Dataset
Taal	Engels

↓ Toon meer

Trefwoorden

higher education relationship management relationship quality
student-faculty relationships confirmatory factor analysis

Thema

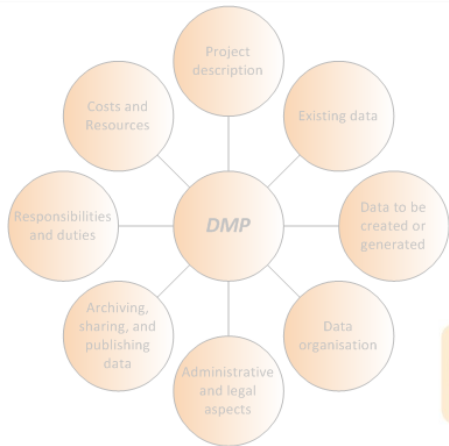
Opvoeding & Onderwijs

En is een pragmatische werkwijze gevonden en gevalideerd

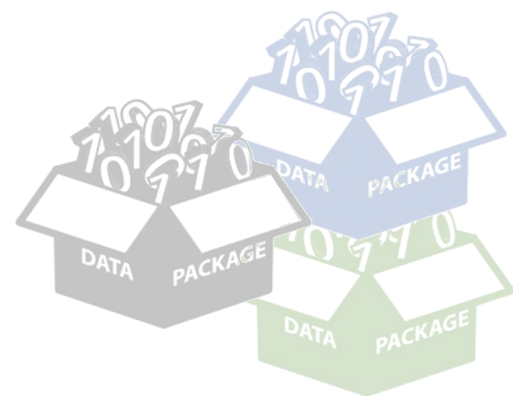
1. DMP (ondersteuning door bibliotheek -> leidt tot projectfolder op Research Drive);
2. Onderzoek uitvoeren door projectteam;
3. Data gereed maken voor publicatie (opschonen, data packages, readme.
Goede instructies: [FAIR Principles - GO FAIR \(go-fair.org\)](https://go-fair.org/);
[Storing and preserving data | Universiteit Utrecht \(uu.nl\)](https://uu.nl/storing-and-preserving-data);
[Policies & Guidelines | 4TU.ResearchData](https://4TU.ResearchData/policies-guidelines));
4. Publiceren data op Sharekit (abstract, metadata, DOI, data);
5. Eventueel ook bijbehorende artikelen op Sharekit publiceren.

Volgende stappen

1. Korte handout/instructie voor de werkwijze
2. Samen met SURF opschalen/uitrollen (daarbij breder kijken dan alleen deze route)
3. Implementatie op HZ
4. (Landelijk) gebruik en feedback monitoren
5. Integratie met Research Drive



SURF RESEARCH DRIVE



SURF SHAREKIT



SURF SHARE
Data Persistent Identifier



HBO kennisbank

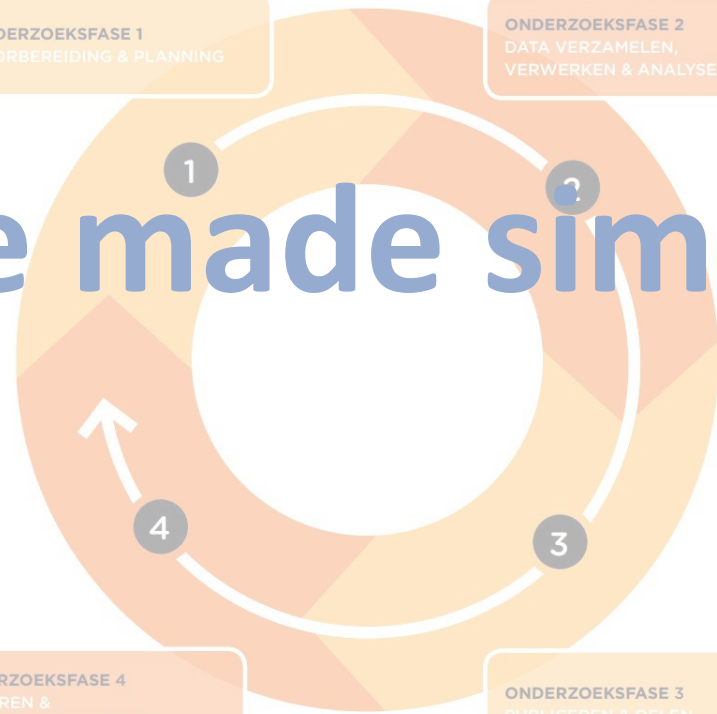
SURF Powered by SURF

ONDERZOEKSFASE 1
VOORBEREIDING & PLANNING

ONDERZOEKSFASE 2
DATA VERZAMELEN,
VERWERKEN & ANALYSEREN

ONDERZOEKSFASE 4
BEWAREN &
HERGEBRUIK VAN DATA

ONDERZOEKSFASE 3
PUBLICEREN & DELEN



“Open science made simple”

Vragen?