



Pitch voorstel: **Onderzoek naar het opzetten van een Nederlands Data Curatie Netwerk**
Namen: **Mijke Jetten, Maaïke Messelink, Tina Reilink, Inge Slouwerhof – Radboud Universiteit**
Datum: **28-01-2019**
Contact details: **e: researchdata@ubn.ru.nl; t: 024 3612863**

1. Doel (probleem definitie, oplossings-georiënteerde aanpak) en afbakening van het onderwerp.

Beschrijving van de pitch & impact op de gemeenschap

Gebaseerd op de principes van Open Science en Wetenschappelijke Integriteit, verwachten universiteiten, journals en subsidiegevers van onderzoekers dat ze hun data voor de lange termijn archiveren (wetenschappelijke integriteit) en, als dit ethisch en juridisch mogelijk is, hun data beschikbaar stellen voor anderen (hergebruik van data).

Dit stelt eisen aan het data curatie proces voor het lange termijn archiveren van onderzoeksdata, wat over het algemeen wordt gedaan door data archieven en de Research Data Management afdelingen van universiteiten en onderzoeksorganisaties (inclusief data scientists, stewards, managers en librarians).

De [FAIR Guiding Principles](#) voor wetenschappelijk data management en datastewardschap bieden basis criteria voor (onder andere) datacuratieprocessen. Eerste stappen zijn dus zeker gezet en de doelen zijn bepaald (FAIR data). Echter, op dit moment hebben we geen duidelijk beeld van het huidige Nederlandse datacuratielandschap. Is het uniform of bestaat er grote variatie in de kwaliteit, structuur, inhoud en context van datacuratie bij de verschillende data archieven, universiteiten en onderzoeksorganisaties?

Gebaseerd op ervaring en informatie aanwezig in de VS, geïnitieerd door het [Data Curation Network Project](#), is het nu de tijd om de krachten in Nederland te bundelen en de mogelijkheid te onderzoeken om een gelijkaardig project (hoewel op kleinere schaal) te lanceren; de Dutch Data Curation Network.

Het idee voor dit project ontstond tijdens de IDCC Barcelona conferentie (Februari 2018), naar aanleiding van de presentatie van het in de VS geïnitieerde project [Data Curation Network Project](#). Ondanks dat de aanpak van het VS project omvangrijk is (een gesubsidieerd vijf jarig project) en zich ook richt op het uitwisselen van personeel, wat waarschijnlijk teveel is voor de Nederlandse versie, was een groep Nederlanders (van universiteiten en data archieven) enthousiast over de ideeën die gepresenteerd werden. Met dit in ons achterhoofd, stellen we ons de volgende vraag: hoe kunnen we onze krachten bundelen en het idee van een Nederlands Data Curatie Network onderzoeken?

De LCRDM taakgroep *Investigating a Digital Curation Network* zal in eerste instantie zijn activiteiten beperken tot een onderzoek naar de bestaande kennis en de wensen en benodigdheden ten aanzien van datacuratie. Afhankelijk van de uitkomst van dit onderzoek, zal de volgende stap het ontwikkelen van een Nederlandse data curatie standaard zijn, binnen een Nederlands Data Curatie Netwerk (project).

2. Waarom is dit probleem instituut overschrijdend?

In Nederland zijn datacuratieprocessen vrij gefragmenteerd georganiseerd en missen we protocollen die een kader vormen voor de curatie van data. Bovendien is er mogelijk een sterke behoefte om een Nederlandse samenwerking op te zetten ten behoeve van datacuratie, vooral omdat we (nog) niet kunnen verwachten dat een universiteit of een onderzoeksorganisatie zelf alle expertise en vaardigheden in dit brede veld van disciplinaire datacuratie bezit. Deze samenwerking wordt daarnaast relevanter als je beseft welk 'retrospectief effect' het kan hebben, namelijk de noodzaak om onderzoekers in een vroeg stadium van het onderzoek voor te bereiden op de data curatie, door middel van het verwerken van de data curatie standaarden in DMP's en RDM trainingen.

3. Deliverable/output, wat is de verwachte opbrengst? (het "product") + ruwe schets.

Deliverable/output

- De behoeften en mogelijkheden onderzoeken rondom het opzetten van een Nederlands Data Curatie Netwerk
- Focus ligt op (1) het in kaart brengen van de huidige kennis en (2) het opzetten van een inventaris van wensen en behoeften betreffende data curatie
- Het in een latere fase, afhankelijk van de uitkomst van dit project, ontwikkelen van data curatie standaarden of zelfs het oppakken van één van de meer uitgebreide doelen van het originele VS project, wellicht met behulp van subsidies.

Deel van het werkplan (nog te besluiten in de taakgroep) kan zijn:

- Inventaris van de bestaande kennis en handelingen, bijvoorbeeld door middel van een questionnaire
- Inventaris van wensen en behoeften, bijvoorbeeld via interviews of focus groepen
- Advies omtrent:
 - De haalbaarheid/wenselijkheid van zo'n netwerk in Nederland, gebaseerd op de uitkomsten van de inventarisaties, inclusief een indicatie van de toegevoegde waarde van een datacuratienetwerk naast bijvoorbeeld het LCRDM of NDE.
 - En, als de uitkomst positief is, een eerste stap naar standaardisatie: de identificatie van wederzijdse kwesties kunnen behandeld worden in zo'n netwerk, en de omvang van data curatie in de volledige data life cycle.

4. Welke expertise/competenties zijn nodig om de taak groep op te zetten? En hoeveel subsidie is nodig?

Welke mensen hebben we nodig?

Iedereen met interesse in data curatie processen, gericht op standaardisatie en/of het vinden van een gemeenschappelijke ideeën, wonend in Nederland. Men kan werkzaam zijn bij een Nederlandse organisatie/universiteit/netwerk of vanuit individuele interesse deel nemen. Het uitgangspunt, echter, is Nederland breed: het zou geweldig zijn als we samen kunnen ontdekken en uiteindelijk iets kunnen ontwikkelen dat bruikbaar is voor alle Nederlandse onderzoeksorganisaties.

De taak groep bestaat idealiter uit vijf of zes enthousiaste mensen om aan het project te werken, en als nodig, een aantal mensen in een adviserende functie.

5. Ingeschatte tijdsduur (max 2 maanden) en planning (gewenste startdatum en datum van afronding) - NB: de hoeveelheid tijd die je per week besteed aan dit onderwerp kan gecoördineerd worden in de taakgroep

De verwachte doorlooptijd is vier maanden (aangezien het proces van het verspreiden en ontvangen van vragenlijsten tijd kost): eind februari – eind mei 2019.

6. Wat weten we al/is al ondernomen en waarom is deze informatie niet toereikend?

- ✓ US Data Connection Network:
<https://sites.google.com/site/datacurationnetwork/home>
- ✓ NPOS – Ambitie 4: Stimuleren Open Science: netwerk experts open science
- ✓ NDE – Strategisch Plan 2019-2020 (beschikbaar 24/9/2018): voortbouwen op NCDD/NDE-plan uit 2016 voor de ontwikkeling van een preservation network in Nederland. (Concept-businessplan uit 2016 – ongepubliceerd; plan bevroren).