Jonas K. Rømer, AU-DCE. jkr@ecos.au.dk . Feb. 2022

# DeIC / DataCite Konsortium: Afrapportering 2021-2022 fra projektet Greenland Ecosystem Monitoring (GEM) og det tilknyttede open data repository under AU's samlede aftale.

## Det strategiske, organisatoriske og økonomiske:

### Hvad er ønsker om og planer for udvikling, herunder politik, funktionalitet, ved såvel eget repositorium som hos DataCite?

Vi planlægger og står over for opstart af en opskalering og modernisering af vores repository. Dette er egenudviklet, og består af en core database, open data website og integrationer + endpoints (DataCite MDS api, Google Dataset Search, Datacite metadata schema 4.3, schema.org/science-on-schema.org)

 - Vores kernedatabase bag repository er fintunet og opdateres med data årligt. Denne vil vi bibeholde da den overordnet er velfungerende, men reviewe og gennemarbejde i mindre grad for at få ryddet op i uhensigtsmæssige dele.

 - Vores website ønsker vi at flytte over på CKAN data portal software (https://ckan.org), for at kunne bruge komponenter fra open source miljøet, øge muligheder for visualisering af data og generelt gøre vores data lettere tilgængelige for flere målgrupper i et system der er lettere at udvikle nye features i, end vores nuværende.

- Modernisering af bagvedliggende teknologier og infrastruktur. Muligvis skift til cloud miljø. Bedre muligheder for lagring af store datamængder som model output og remote sensing data.

 - Vores integrationer til datacite er på plads og alle datasets registreres med DOI. Vi vil forbedre mulighederne for versionering af datasets i vores database, sende opdaterede versionsnumre til datacite.

I forbindelse med de initiativer vi vil iværksætte i år, planlægger vi en mere open source tilgang også til vores repository kode og tilpasninger, dokumentation mv. Vi vil oprette et github projekt og tilhørende github pages, så alle - både deltagende parter i GEM projektet men også andre interesserede, kan følge med og bidrage.

### Hvorledes arbejdes der for udbredelse af DOI brug, samt eventuelt andre PID løsninger?

På to årlige koordinationsgruppemøder i projektet inddrages alle databidragere i udviklingen af DOI brug, metadatastandarder mv. Ved download af vores data, får brugere beskrivelser med, som angiver hvordan data skal citeres med DOI, og link til hvordan man kan se datasættet registreret hos DataCite (hos datacite kan man også hente reference i forskellige bibliografiske formater).

Med datacite konsortiet og aftalen på plads, er vi naturligvis opmærksomme på andre projekter hvor lignende systemer og DOI registrering kan anvendes. Ligeledes afventer vi implementeringen af ERDA systemet på AU generelt, og de muligheder dette skal give alle for opbevaring af data og registrering af DOIs.

### Hvilke problemer løser DOI’er, og hvilke løser de ikke?

Første praktiske fordel de giver os, er en unik nøgle på vores datasæt. Det frigør os fra tidligere tvivl om hvilket datasæt der var tale om, og problemer med at matche ved opdateringer af databasen hvis navne med mere blev ændret.

Aktuelt har vi problemer omkring hvordan og hvor vores metadata vedligeholdes: databidragerne har tidligere vedligeholdt metadata decentralt og leveret dette sammen med data. Deres dokumenter er dog ikke opdateret med DOIs som vi så mangler på datasets ved import. Derfor må vi etablere en ny central vedligeholdelse af metadata for at sikre konsistent og bedre arbejdsgange omkring opdatering og vedligehold af metadata.

Vi har endnu ikke oplevet øgede henvisninger, citationer eller nye brugere af vores data i væsentlig grad, på trods af registrering i datacites søgesystemer, bedre synlighed på google, DOIs.

### Hvilken betydning har den nationale FAIR strategi?

Den er selvfølgelig altafgørende, både for udviklingen af open science og open data, men også for at vi får nogle praktiske værktøjer og midler som muliggør registrering af DOIs, langtidsopbevaring af store datasæts, publicering af data.

## Det operative:

### Hvilke gode/dårlige erfaringer han man gjort sig med DataCite?

 + Grundig og letforståelig dokumentation.

 + Engageret team.

 + Velfungerende services til registrering af DOIs.

 + interessante workshops / webinars (kan ikke huske hvad de kalder de events de holder).

 - vedr. workshop / webinars og de events de holder, samt mails de udsender om dem. Er de kun åbne for medlemmer? Alle interesserede bør kunne signe up til nyhedsbrev på datacites hjemmeside og deltage i onlinemøderne.

### Er der konkrete ønsker til funktionalitet, DataCite Fabrica, DataCite MDS API?

 - Ikke umiddelbart. Har et konkret problem med utf-8 og danske tegn via MDS api, men kan lige så vel være en fejl i vores kode. Jeg har endnu ikke kontaktet dem om problemet, så regner med det kan løses ved at kontakte deres support og afprøve lidt mere.